



US009677110B2

(12) **United States Patent**  
**Cunningham**

(10) **Patent No.:** **US 9,677,110 B2**  
(45) **Date of Patent:** **\*Jun. 13, 2017**

(54) **METHODS FOR THE IDENTIFICATION OF ANTIBIOTICS THAT ARE NOT SUSCEPTIBLE TO ANTIBIOTIC RESISTANCE IN PSEUDOMONAS AERUGINOSA**

5,981,280 A 11/1999 Fang et al.  
7,045,337 B2 5/2006 Schultz et al.  
7,081,341 B2 7/2006 Cunningham

FOREIGN PATENT DOCUMENTS

(71) Applicant: **Wayne State University**, Detroit, MI (US)

WO 96/06106 A1 2/1996  
WO 00/32619 A1 6/2000  
WO 01/42445 A2 6/2001  
WO 02/29019 A2 4/2002  
WO 03/029459 A2 4/2003  
WO 2004/003511 A2 1/2004  
WO 2006/115570 A2 11/2006

(72) Inventor: **Philip R. Cunningham**, Troy, MI (US)

(73) Assignee: **Wayne State University**, Detroit, MI (US)

OTHER PUBLICATIONS

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

Tapprich, W. E., et al., "Involvement of Bases 787-795 of *Escherichia coli* 16S Ribosomal RNA in Ribosomal Subunit Association," Proceedings of the National Academy of Sciences of the United States of America, Feb. 1986, pp. 556-560, vol. 83, No. 3.  
Tapprich, W. E., et al., "Mutation at Position 791 in *Escherichia coli* 16S Ribosomal RNA Affects Processes Involved in the Initiation of Protein Synthesis," Proceedings of the National Academy of Sciences of the United States of America, Jul. 1989, pp. 4927-4931, vol. 86, No. 13.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **14/746,043**

(22) Filed: **Jun. 22, 2015**

(65) **Prior Publication Data**

US 2015/0284765 A1 Oct. 8, 2015

**Related U.S. Application Data**

(62) Division of application No. 13/654,911, filed on Oct. 18, 2012, now Pat. No. 9,062,339, which is a division of application No. 11/914,062, filed as application No. PCT/US2006/018187 on May 11, 2006, now Pat. No. 8,293,517.

Triman, K., et al., Isolation of Temperature-Sensitive Mutants of 16 S rRNA in *Escherichia coli*, Journal of Molecular Biology, Oct. 1989, pp. 645-653, vol. 209, No. 4.

Tsiodras, S., et al., "Linezolid Resistance in a Clinical Isolate of *Staphylococcus aureus*," Lancet, Jul. 2001, pp. 207-208, vol. 358, No. 9277.

(60) Provisional application No. 60/680,134, filed on May 11, 2005.

Vicens, Q., et al., "Crystal Structure of Geneticin Bound to a Bacterial 16S Ribosomal RNA a Site Oligonucleotide," Journal of Molecular Biology, Feb. 2003, pp. 1175-1188, vol. 326, No. 4.

Vicens, Q., et al., "Crystal Structure of Paromomycin Docked into the Eubacterial Ribosomal Decoding a Site," Structure, Aug. 2001, pp. 647-658, vol. 9, No. 8.

(51) **Int. Cl.**

**C12Q 1/18** (2006.01)  
**C12N 15/10** (2006.01)  
**C12N 15/67** (2006.01)  
**C12N 15/70** (2006.01)  
**C12Q 1/02** (2006.01)  
**C12Q 1/10** (2006.01)  
**C12Q 1/68** (2006.01)

Vila-Sanjurjo, A., et al., "Isolation of Kasugamycin Resistant Mutants in the 16 S Ribosomal RNA of *Escherichia coli*," Journal of Molecular Biology, Oct. 1999, pp. 1-8, vol. 293, No. 1.

Vila-Sanjurjo, A., et al., "Mutational Analysis of the Conserved Bases C1402 and A1500 in the Center of the Decoding Domain of *Escherichia coli* 16 S rRNA Reveals an Important Tertiary Interaction," Journal of Molecular Biology, May 2001, pp. 457-463, vol. 308, No. 3.

(52) **U.S. Cl.**

CPC ..... **C12Q 1/18** (2013.01); **C12N 15/1034** (2013.01); **C12N 15/1086** (2013.01); **C12N 15/67** (2013.01); **C12N 15/70** (2013.01); **C12Q 1/025** (2013.01); **C12Q 1/10** (2013.01); **C12Q 1/6876** (2013.01); **C12Q 2600/156** (2013.01); **C12Q 2600/158** (2013.01); **G01N 2333/21** (2013.01); **G01N 2500/10** (2013.01)

Voulgaris, J., et al., "Increased rrn Gene Dosage Causes Intermittent Transcription of rRNA in *Escherichia coli*," Journal of Bacteriology, Jul. 1999, pp. 4170-4175, vol. 181, No. 14.

(Continued)

(58) **Field of Classification Search**

None  
See application file for complete search history.

Primary Examiner — Nancy Treptow

(74) Attorney, Agent, or Firm — Senniger Power LLP

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,772,555 A 9/1988 De Boer  
4,873,316 A 10/1989 Meade et al.  
5,821,058 A 10/1998 Smith et al.

(57) **ABSTRACT**

The "instant evolution" system was initially developed in *E. coli*, primarily because of the ease with which this organism can be genetically manipulated. Because many of the functionally important regions of rRNA are conserved among bacteria, drug leads developed against conserved targets in the *E. coli* system may produce broad-spectrum anti-infectives. In order to develop a system to produce narrow-spectrum anti-infectives, herein we disclose methods for identifying functional mutant *P. aeruginosa* ribosomes suitable as drug targets and for identifying drug candidates that do not bind to the human 16S rRNA.

**18 Claims, 53 Drawing Sheets**

(56)

## References Cited

## OTHER PUBLICATIONS

- Wada, K., et al., "Codon Usage Tabulated From the GenBank Genetic Sequence Data," *Nucleic Acids Research*, May 1992, pp. 2111-2118, vol. 20 Supplement.
- Wallace, C. K., et al., "Optimal Antibiotic Therapy in Cholera," *Bulletin of the World Health Organization*, 1968, pp. 239-245, vol. 39, No. 2.
- Wang, Y., et al., "Specificity of Aminoglycoside Binding to RNA Constructs Derived from the 16S rRNA Decoding Region and the HIV-RRE Activator Region," *Biochemistry*, Jan. 1997, pp. 768-779, vol. 36, No. 4.
- Wanke, C. A., "To Know *Escherichia coli* is to Know Bacterial Diarrheal Disease," *Clinical Infectious Diseases*, 2001, pp. 1710-1712, vol. 32, No. 12.
- Wilson, K. S., et al., "Mapping the Position of Translational Elongation Factor EF-G in the Ribosome by Directed Hydroxyl Radical Probing," *Cell*, Jan. 1998, pp. 131-139, vol. 92, No. 1.
- Wimberly, B. T., et al., "Structure of the 30S Ribosomal Subunit," *Nature*, Sep. 2000, pp. 327-339, vol. 407, No. 6802.
- Winkler, F. K., et al., "Structure-Based Approaches in Modern Drug Discovery Research," Ernst Schering Research Foundation Workshop, 2001, pp. 123-142, No. 34.
- Winoto, A., et al., "A Novel, Inducible and T Cell-Specific Enhancer Located at the 3' End of the T Cell Receptor Alpha Locus," *The EMBO Journal*, Mar. 1989, pp. 729-733, vol. 8, No. 3.
- Xiong, L., et al., "Oxazolidinone Resistance Mutations in 23S rRNA of *Escherichia coli* Reveal the Central Region of Domain V as the Primary Site of Drug Action," *Journal of Bacteriology*, Oct. 2000, pp. 5325-5331, vol. 182, No. 19.
- Yanisch-Perron, C., et al., "Improved M13 Phage Cloning Vectors and Host Strains: Nucleotide Sequences of the M13mp18 and pUC19 Vectors," *Gene*, 1985, pp. 103-119, vol. 33, No. 1.
- Yoshizawa, S., et al., "Structural Origins of Gentamicin Antibiotic Action," *The EMBO Journal*, Nov. 1998, pp. 6437-6448, vol. 17, No. 22.
- Yusupov, M. M., et al., "Crystal Structure of the Ribosome at 5.5 Å Resolution," *Science*, 2001, pp. 883-896, vol. 292, No. 5518.
- Yusupova, G. Z., et al., "The Path of Messenger RNA Through the Ribosome," *Cell*, Jul. 2001, pp. 233-241, vol. 106, No. 2.
- Zhang, K., et al., "Assessing Reliability of Gene Clusters from Gene Expression Data," *Functional & Integrative Genomics*, Nov. 2000, pp. 156-173, vol. 1, No. 3.
- Zhang, X., et al., "Quinolone Antibiotics Induce Shiga Toxin-Encoding Bacteriophages, Toxin Production, and Death in Mice," *The Journal of Infectious Diseases*, Feb. 2000, pp. 664-670, vol. 181, No. 2.
- Supplemental European Search Report for EP 03 76 2329 dated Nov. 10, 2005.
- International Search Report for PCT/US2006/018320 dated Nov. 28, 2007.
- Supplemental European Search Report for EP 06 75 9609 dated Jan. 13, 2009.
- Evans, G. A., "The Oxazolidinones," *Current Infectious Disease Reports*, Feb. 2002, pp. 17-27, vol. 4, No. 1.
- Fauci, A. S., et al., "Emerging Infectious Diseases: A 10-Year Perspective from the National Institute of Allergy and Infectious Diseases," *Emerging Infectious Diseases*, Apr. 2005, pp. 519-525, vol. 11, No. 4.
- Fourmy, D., et al., "Structure of the a Site of *Escherichia coli* 16S Ribosomal RNA Complexed with an Aminoglycoside Antibiotic," *Science*, Nov. 1996, pp. 1367-1371, vol. 274, No. 5291.
- Gabashvili, I. S., et al., "Major Rearrangements in the 70S Ribosomal 3D Structure Caused by a Conformational Switch in 16S Ribosomal RNA," *The EMBO Journal*, Nov. 1999, pp. 6501-6507, vol. 18, No. 22.
- Goeddel, D. V., "Systems for Heterologous Gene Expression," *Methods in Enzymology*, 1990, pp. 3-7, vol. 185.
- Gomi, H., et al., "In Vitro Antimicrobial Susceptibility Testing of Bacterial Enteropathogens Causing Traveler's Diarrhea in Four Geographic Regions," *Antimicrobial Agents and Chemotherapy*, Jan. 2001, pp. 212-216, vol. 45, No. 1.
- Gonzales, R. D., et al., "Infections Due to Vancomycin-Resistant *Enterococcus faecium* Resistant to Linezolid," *Lancet*, Apr. 2001, p. 1179, vol. 357, No. 9263.
- Gottesman, S., "Minimizing Proteolysis in *Escherichia coli*: Genetic Solutions," *Methods in Enzymology*, 1990, pp. 119-129, vol. 185.
- Govantes, F., et al., "Mechanism of Translational Coupling in the nifLA Operon of *Klebsiella pneumoniae*," *The EMBO Journal*, Apr. 1998, pp. 2368-2377, vol. 17, No. 8.
- Guerrant, R. L., et al., "Magnitude and Impact of Diarrheal Diseases," *Archives of Medical Research*, Jul.-Aug. 2002, pp. 351-355, vol. 33, No. 4.
- Gutell, R. R., "Collection of Small Subunit (16S- and 16S-Like) Ribosomal RNA Structures: 1994," *Nucleic Acids Research*, Sep. 1994, pp. 3502-3507, vol. 22, No. 17.
- Gutell, R. R., et al., "Identifying Constraints on the Higher-Order Structure of RNA: Continued Development and Application of Comparative Sequence Analysis Methods," *Nucleic Acids Research*, Nov. 1992, pp. 5785-5795, vol. 20, No. 21.
- Haddad, J., et al., "Design of Novel Antibiotics That Bind to the Ribosomal Acyltransfer Site," *Journal of the American Chemical Society*, Apr. 2002, pp. 3229-3237, vol. 124, No. 13.
- Hanahan, D., "Studies on Transformation of *Escherichia coli* With Plasmids," *Journal of Molecular Biology*, Jun. 1983, pp. 557-580, vol. 116, No. 4.
- Hancock, R. E., et al., "Peptide Antibiotics," *Antimicrobial Agents and Chemotherapy*, Jun. 1999, pp. 1317-1323, vol. 43, No. 6.
- Harms, J., et al., "High Resolution Structure of the Large Ribosomal Subunit from a Mesophilic Eubacterium," *Cell*, Nov. 2001, pp. 679-688, vol. 107, No. 5.
- Hartman, A. B., et al., "Epidemiology of Tetracycline Resistance Determinants in *Shigella* spp. and Enteroinvasive *Escherichia coli*: Characterization and Dissemination of tet(A)-1," *Journal of Clinical Microbiology*, Mar. 2003, pp. 1023-1032, vol. 41, No. 3.
- Herr, W., et al., "Mechanism of Ribosomal Subunit Association: Discrimination of Specific Sites in 16 S RNA Essential for Association Activity," *Journal of Molecular Biology*, Jun. 1979, pp. 433-449, vol. 130, No. 4.
- Herrero, I. A., et al., "Nosocomial Spread of Linezolid-Resistant, Vancomycin-Resistant *Enterococcus faecium*," *The New England Journal of Medicine*, Mar. 2002, pp. 867-869, vol. 346, No. 11.
- Higuchi, R., "Using PCR to Engineer DNA," *PCR Technology* (Erlich, H.A., ed.), 1989, pp. 61-70.
- Hui, A., et al., "Directing Ribosomes to a Single mRNA Species: A Method to Study Ribosomal RNA Mutations and Their Effects on Translation of a Single Messenger in *Escherichia coli*," *Methods in Enzymology*, 1987, pp. 432-452, vol. 153.
- Hui, A., et al., "Specialized Ribosome System: Preferential Translation of a Single mRNA Species by a Subpopulation of Mutated Ribosomes in *Escherichia coli*," *Proceedings of the National Academy of Sciences of the United States of America*, Jul. 1987, pp. 4762-4766, vol. 84, No. 14.
- Hwang, S., et al., "Inhibition of Gene Expression in Human Cells Through Small Molecule-RNA Interactions," *Proceedings of the National Academy of Sciences of the United States of America*, Nov. 1999, pp. 12997-13002, vol. 96, No. 23.
- John, T. J., "Emerging & Re-Emerging Bacterial Pathogens in India," *The Indian Journal of Medical Research*, Jan. 1996, pp. 4-18, vol. 103.
- Jones, R. N., et al., "Linezolid-Resistant *Enterococcus faecium* Isolated from a Patient Without Prior Exposure to an Oxazolidinone: Report from the Sentry Antimicrobial Surveillance Program," *Diagnostic Microbiology and Infectious Disease*, Feb. 2002, pp. 137-139, vol. 42, No. 2.
- Kaufman, R. J., et al., "Translational Efficiency of Polycistronic mRNAs and Their Utilization to Express Heterologous Genes in Mammalian Cells," *The EMBO Journal*, Jan. 1987, pp. 187-193, vol. 6, No. 1.

(56)

## References Cited

## OTHER PUBLICATIONS

- Keren-Zur, M., et al., "Localization of the Decoding Region on the 30S *Escherichia coli* Ribosomal Subunit by Affinity Immunoelectron Microscopy," Proceedings of the National Academy of Sciences of the United States of America, Mar. 1979, pp. 1054-1058, vol. 76, No. 3.
- Kessel, M., et al., "Murine Developmental Control Genes," Science, Jul. 1990, pp. 374-379, vol. 249, No. 4967.
- Khan, W. A., et al., "Randomised Controlled Comparison of Single-Dose Ciprofloxacin and Doxycycline for Cholera Caused by *Vibrio cholerae* 01 or 0139," Lancet, Aug. 1996, pp. 296-300, vol. 348, No. 9023.
- Kieber, Emmons, T., et al., "Therapeutic Peptides and Peptidomimetics," Current Opinion in Biotechnology, Aug. 1997, pp. 435-441, vol. 8, No. 4.
- Kloss, P., et al., "Resistance Mutations in 23 S rRNA Identify the Site of Action of the Protein Synthesis Inhibitor Linezolid in the Ribosomal Peptidyl Transferase Center," Journal of Molecular Biology, Nov. 1999, pp. 93-101, vol. 294, No. 1.
- Klostermeier, D., et al., "A Three-Fluorophore FRET Assay for High-Throughput Screening of Small-Molecule Inhibitors of Ribosome Assembly," Nucleic Acids Research, 2004, pp. 2707-2715, vol. 32, No. 9.
- Koosha, H., et al., "Alterations in the Peptidyltransferase and Decoding Domains of Ribosomal RNA Suppress Mutations in the Elongation Factor G Gene," RNA, Aug. 2000, pp. 1166-1173, vol. 6, No. 8.
- Kozak, M., "Regulation of Translation via mRNA Structure in Prokaryotes and Eukaryotes," Gene, 2005, pp. 13-37, vol. 361.
- Krzyzosiak, W., et al., "In Vitro Synthesis of 16S Ribosomal RNA Containing Single Base Changes and Assembly into a Functional 30S Ribosome," Biochemistry, Apr. 1987, pp. 2353-2364, vol. 26, No. 8.
- Kurjan, J., et al., "Structure of a Yeast Pheromone Gene (MF alpha): A Putative Alpha-Factor Precursor Contains Four Tandem Copies of Mature Alpha-Factor," Cell, Oct. 1982, pp. 933-943, vol. 30, No. 3.
- Laios, E., et al., "Combinatorial Genetic Technology for the Development of New Anti-Infectives," Archives of Pathology & Laboratory Medicine, Dec. 2004, pp. 1351-1359, vol. 128, No. 12.
- Lam, K. S., et al., "Application of a Dual Color Detection Scheme in the Screening of a Random Combinatorial Peptide Library," Journal of Immunological Methods, Mar. 1995, pp. 219-223, vol. 180, No. 2.
- Lee, K., et al., "Genetic Approaches to Studying Protein Synthesis: Effects of Mutations at Psi516 and A535 in *Escherichia coli* 16S rRNA," Symposium: Translational Control: A Mechanistic Perspective at the Experimental Biology 2001 Meeting.
- Lee, K., et al., "Genetic Approaches to Studying Protein Synthesis: Effects of Mutations at Psi516 and A535 in *Escherichia coli* 16S rRNA," The Journal of Nutrition, Nov. 2001, pp. 2994S-3004S, vol. 131, No. 11.
- Lee, K., et al., "In Vivo Determination of RNA Structure-Function Relationships: Analysis of the 790 Loop in Ribosomal RNA," Journal of Molecular Biology, Jun. 1997, pp. 732-743, vol. 269, No. 5.
- Lee, K., et al., "Genetic Analysis of the Shine-Dalgarno Interaction: Selection of Alternative Functional mRNA-rRNA Combinations," RNA, Dec. 1996, pp. 1270-1285, vol. 2, No. 12.
- Lesley, S. A., et al., "Use of in Vitro Protein Synthesis from Polymerase Chain Reaction-Generated Templates to Study Interaction of *Escherichia coli* Transcription Factors with Core RNA Polymerase and for Epitope Mapping of Monoclonal Antibodies," The Journal of Biological Chemistry, Feb. 1991, pp. 2632-2638, vol. 266, No. 4.
- Levy, S. B., et al., "Antibacterial Resistance Worldwide: Causes, Challenges and Responses," Nature Medicine, Dec. 2004, pp. 5122-5129, vol. 10, Supplement 12.
- Levy, S. B., "Antibiotic Resistance: Consequences of Inaction," Clinical Infectious Diseases, Sep. 2001, pp. 5124-5129, vol. 33, Supplement 3.
- Lin-Goerke, J. L., et al., "PCR-Based Random Mutagenesis Using Manganese and Reduced dNTP Concentration," BioTechniques, Sep. 1997, pp. 409-412, vol. 23, No. 3.
- Lind, K. E., et al., "Structure-Based Computational Database Screening, In Vitro Assay, and NMR Assessment of Compounds that Target TAR RNA," Chemistry & Biology, Feb. 2002, pp. 185-193, vol. 9, No. 2.
- Lindenbaum, J., et al., "Antibiotic Therapy of Cholera in Children," Bulletin of the World Health Organization, 1967, pp. 529-538, vol. 37, No. 4.
- Llano-Sotelo, B., et al., "Aminoglycosides Modified by Resistance Enzymes Display Diminished Binding to the Bacterial Ribosomal Aminoacyl-tRNA Site," Chemistry & Biology, Apr. 2002, pp. 455-463, vol. 9, No. 4.
- Lodmell, J. S., et al., "A Conformational Switch in *Escherichia coli* 16S Ribosomal RNA During Decoding of Messenger RNA," Science, Aug. 1997, pp. 1262-1267, vol. 277, No. 5330.
- Adachi, J. A., et al., "Natural History of Enterotoxigenic and Enterotoxigenic *Escherichia coli* Infection Among US Travelers to Guadalajara, Mexico," The Journal of Infectious Diseases, 2002, pp. 1681-1683, vol. 185, No. 11.
- Adang, A. E. P., et al., "Case Histories of Peptidomimetics: Progression from Peptides to Drugs," Rescued des Travaux Chimiques des Pays-Bas, 1994, pp. 63-78, vol. 113.
- Agalarov, S. C., et al., "Structure of the S15,S6,S18-rRNA Complex: Assembly of the 30S Ribosome Central Domain," Science, Apr. 2000, pp. 107-113, vol. 288, No. 5463.
- Ahn, J. M., et al., "Peptidomimetics and Peptide Backbone Modifications," Mini Reviews in Medicinal Chemistry, Oct. 2002, pp. 463-473, vol. 2, No. 5.
- Andersson, D. I., et al., "Antibiotic Resistance Here to Stay? Compensatory Mutations Restore Virulence of Resistant Bacteria," Lakartidningen, 1998, pp. 3943-3944, vol. 95, No. 3940.
- Ang, J. Y., et al., "Antibacterial Resistance," Indian Journal of Pediatrics, Mar. 2004, pp. 229-239, vol. 71, No. 3.
- Aoki, H., et al., "Oxazolidinone Antibiotics target the P Site on *Escherichia coli* Ribosomes," Antimicrobial Agents and Chemotherapy, Apr. 2002, pp. 1080-1085, vol. 46, No. 4.
- Asai, T., et al., "Construction and Initial Characterization of *Escherichia coli* Strains With Few or No Intact Chromosomal rRNA Operons," Journal of Bacteriology, Jun. 1999, pp. 3803-3809, vol. 181, No. 12.
- Baldari, C., et al., "A Novel Leader Peptide Which Allows Efficient Secretion of a Fragment of Human Interleukin 1 beta in *Saccharomyces cerevisiae*," The EMBO Journal, Jan. 1987, pp. 229-234, vol. 6, No. 1.
- Banerji, J., et al., "A Lymphocyte-Specific Cellular Enhancer is Located Downstream of the Joining Region in Immunoglobulin Heavy Chain Genes," Cell, Jul. 1983, pp. 729-740, vol. 33, No. 3.
- Barrick, J. E., et al., "Selection of RNA-Binding Peptides Using mRNA-Peptide Fusions," Methods, Mar. 2001, pp. 287-293, vol. 23, No. 3.
- Barrick, J. E., et al., "Sequence Analysis of an Artificial Family of RNA-Binding Peptides," Protein Science, Nov. 2002, pp. 2688-2696, vol. 11, No. 11.
- Batey, R. T., et al., "Interaction of the *Bacillus stearothermophilus* Ribosomal Protein S15 with 16 S rRNA: II. Specificity Determinants of RNA-Protein Recognition," Journal of Molecular Biology, Aug. 1996, pp. 550-567, vol. 261, No. 4.
- Batey, R. T., et al., "Preparation of Isotopically Labeled Ribonucleotides for Multidimensional NMR Spectroscopy of RNA," Nucleic Acids Research, Sep. 1992, pp. 4515-4523, vol. 20, No. 17.
- Bhattacharya, S. K., et al., "Multidrug-Resistant *Shigella dysenteriae* Type 1 in South Asia," The Lancet Infectious Diseases, Dec. 2003, p. 755, vol. 3, No. 12.
- Bjorkman, J., et al., "Effects of Environment on Compensatory Mutations to Ameliorate Costs of Antibiotic Resistance," Science, Feb. 2000, pp. 1479-1482, vol. 287, No. 5457.

(56)

## References Cited

## OTHER PUBLICATIONS

- Bodhidatta, L., et al., "Bacterial Enteric Pathogens in Children With Acute Dysentery in Thailand: Increasing Importance of Quinolone-Resistant *Campylobacter*," *The Southeast Asian Journal of Tropical Medicine and Public Health*, Dec. 2002, pp. 752-757, vol. 33, No. 4.
- Bottger, E. C., "Resistance to Drugs Targeting Protein Synthesis in *Mycobacteria*," *Trends in Microbiology*, Oct. 1994, pp. 416-421, vol. 2, No. 10.
- Brodersen, D. E., et al., "Crystal Structure of the 30 S Ribosomal Subunit From *Thermus thermophilus*: Structure of the Proteins and Their Interactions with 16 S RNA," *Journal of Molecular Biology*, Feb. 2002, pp. 725-768, vol. 316, No. 3.
- Brosius, J., et al., "Construction and Fine Mapping of Recombinant Plasmids Containing the *rnb* Ribosomal RNA Operon of *E. coli*," *Plasmid*, Jul. 1981, pp. 112-118, vol. 6, No. 1.
- Brow, D. A., et al., "Protection of Ribosomal RNA from Kethoxal in Polyribosomes. Implication of Specific Sites in Ribosome Function," *Journal of Molecular Biology*, Jan. 1983, pp. 27-46, vol. 163, No. 1.
- Bursavich, M. G., et al., "Designing Non-Peptide Peptidomimetics in the 21st Century: Inhibitors Targeting Conformational Ensembles," *Journal of Medicinal Chemistry*, Jan. 2002, pp. 541-558, vol. 45, No. 3.
- Byrne, G. W., et al., "Multiplex Gene Regulation: A Two-Tiered Approach to Transgene Regulation in Transgenic Mice," *Proceedings of the National Academy of Sciences of the United States of America*, Jul. 1989, pp. 5473-5477, vol. 86, No. 14.
- Calame, K., et al., "Transcriptional Controlling Elements in the Immunoglobulin and T Cell Receptor Loci," *Advances in Immunology*, 1988, pp. 235-275, vol. 43.
- Calos, M. P., "DNA Sequence for a Low-Level Promoter of the Lac Repressor Gene and an 'Up' Promoter Mutation," *Nature*, Aug. 1978, pp. 762-765, vol. 274, No. 5673.
- Cannone, J. J., et al., "The Comparative RNA Web (CRW) Site: An Online Database of Comparative Sequence and Structure Information for Ribosomal, Intron, and Other RNAs," *BMC Bioinformatics*, Jul. 2002, p. 15, vol. 3, No. 1.
- Capaldi, D. C., et al., "Use of the 1-(2-fluorophenyl)-4-methoxypiperidin-4-yl (FPMP) and Related Protecting Groups in Oligoribonucleotide Synthesis: Stability of Internucleotide Linkages to Aqueous Acid," *Nucleic Acids Research*, Jun. 1994, pp. 2209-2216, vol. 22, No. 12.
- Carter, A. P., et al., "Crystal Structure of an Initiation Factor Bound to the 30S Ribosomal Subunit," *Science*, Jan. 2001, pp. 498-501, vol. 291, No. 5503.
- Carter, A. P., et al., "Functional Insights From the Structure of the 30S Ribosomal Subunit and Its Interactions With Antibiotics," *Nature*, Sep. 2000, pp. 340-348, vol. 407, No. 6802.
- Carter-Muenchau, P., et al., "Growth-Rate-Dependent Regulation of 6-phosphogluconate Dehydrogenase Level Mediated by an Anti-Shine-Dalgarno Sequence Located Within the *Escherichia coli* *gnd* Structural Gene," *Proceedings of the National Academy of Sciences of the United States of America*, Feb. 1989, pp. 1138-1142, vol. 86, No. 4.
- Castagnoli, L., et al., "Alternative Bacteriophage Display Systems," *Combinatorial Chemistry & High Throughput Screening*, Apr. 2001, pp. 121-133, vol. 4, No. 2.
- Cha, J., et al., "New Vectors for Direct Cloning of PCR Products," *Gene*, Dec. 1993, pp. 369-370, vol. 136, Nos. 1-2.
- Chen, H., et al., "Determination of the Optimal Aligned Spacing Between the Shine-Dalgarno Sequence and the Translation Initiation Codon of *Escherichia coli* mRNAs," *Nucleic Acids Research*, Nov. 1994, pp. 4953-4957, vol. 22, No. 23.
- Chopra, I., et al., "The Role of Mutators in the Emergence of Antibiotic-Resistant Bacteria," *Drug Resistant Updates*, Jun. 2003, pp. 137-145, vol. 6, No. 3.
- Chow, C. S., et al., "A Structural Basis for RNA-Ligand Interactions," *Chemical Reviews*, Aug. 1997, pp. 1489-1514, vol. 97, No. 5.
- Clarke, S. C., et al., "Enteropathogenic *Escherichia coli* Infection: History and Clinical Aspects," *British Journal of Biomedical Science*, 2002, pp. 123-127, vol. 59, No. 2.
- Clarke, S. C., et al., "Virulence of Enteropathogenic *Escherichia coli*, A Global Pathogen," *Clinical Microbiology Reviews*, Jul. 2003, pp. 365-378, vol. 16, No. 3.
- Collins, M., et al., "Methicillin-Resistant *Staphylococcus aureus* (MRSA) in the Practice of Otolaryngology—An Emerging Community Acquired Organism?," *Current Opinion in Otolaryngology & Head and Neck Surgery*, Jun. 2003, pp. 179-183, vol. 11, No. 3.
- Cornelis, P., et al., "Cloning and Analysis of the Gene for the Major Outer Membrane Lipoprotein from *Pseudomonas aeruginosa*," *Molecular Microbiology*, Mar. 1989, pp. 421-428, vol. 3, No. 3.
- Cunningham, P. R., et al., "Functional Effects of Base Changes Which Further Define the Decoding Center of *Escherichia coli* 16S Ribosomal RNA: Mutation of C1404, G1405, C1496, G1497, and U1498," *Biochemistry*, Jul. 1993, pp. 7172-7180, vol. 32, No. 28.
- Cunningham, P. R., et al., "The Role of 16S RNA in Ribosome Function: Single Base Alterations and Their Effect on in Vitro Protein Synthesis," *Archivos de Biología y Medicina Experimentales*, Dec. 1988, pp. 393-401, vol. 21, Nos. 3-4.
- Danner, S., et al., "T7 Phage Display: A Novel Genetic Selection System for Cloning RNA-Binding Proteins From cDNA Libraries," *Proceedings of the National Academy of Sciences of the United States of America*, 2001, pp. 12954-12959, vol. 98, No. 23.
- de Boer, H. A., et al., "The *tac* Promoter: A Functional Hybrid Derived from the *trp* and *lac* Promoters," *Proceedings of the National Academy of Sciences of the United States of America*, Jan. 1983, pp. 21-25, vol. 80, No. 1.
- de Stasio, E. A., et al., "Mutations in 16S Ribosomal RNA Disrupt Antibiotic—RNA Interactions," *The EMBO Journal*, Apr. 1989, pp. 1213-1216, vol. 8, No. 4.
- Depardieu, F., et al., "VanD-Type Vancomycin-Resistant *Enterococcus faecium* 10/96A," *Antimicrobial Agents and Chemotherapy*, Jan. 2003, pp. 7-18, vol. 47, No. 1.
- Denman, R., et al., "In Vitro Assembly of 30S and 70S Bacterial Ribosomes from 16S RNA Containing Single Base Substitutions, Insertions, and Deletions Around the Decoding Site (C1400)," *Biochemistry*, Feb. 1989, pp. 1002-1011, vol. 28, No. 3.
- Dessen, A., et al., "Molecular Mechanisms of Antibiotic Resistance in Gram-Positive Pathogens," *Current Drug Targets. Infectious Disorders*, May 2001, pp. 63-77, vol. 1, No. 1.
- Dower, W. J., et al., "High Efficiency Transformation of *E. coli* by High Voltage Electroporation," *Nucleic Acids Research*, Jul. 1988, pp. 6127-6145, vol. 16, No. 13.
- Edlund, T., et al., "Cell-Specific Expression of the Rat Insulin Gene: Evidence for Role of Two Distinct 5' Flanking Elements," *Science*, Nov. 1985, pp. 912-916, vol. 230, No. 4728.
- Enright, M. C., "The Evolution of a Resistant Pathogen—The Case of MRSA," *Current Opinion in Pharmacology*, Oct. 2003, pp. 474-479, vol. 3, No. 5.
- Luria, S. E., et al., "Hybridization Between *Escherichia coli* and *Shigella*," *Journal of Bacteriology*, Oct. 1957, pp. 461-476, vol. 74, No. 4.
- Lynch, S. R., et al., "Comparison of X-Ray Crystal Structure of the 30S Subunit-Antibiotic Complex with NMR Structure of Decoding Site Oligonucleotide-Paromomycin Complex," *Structure*, Jan. 2003, pp. 43-53, vol. 11, No. 1.
- Maden, B. E., "The Numerous Modified Nucleotides in Eukaryotic Ribosomal RNA," *Proceedings of the National Academy of Sciences of the United States of America*, 1990, pp. 241-303, vol. 39.
- Magnet, S., et al., "Molecular Insights into Aminoglycoside Action and Resistance," *Chemical Reviews*, Feb. 2005, pp. 477-498, vol. 105, No. 2.
- Maidak, B. L., et al., "The Ribosomal Database Project (RDP)," *Nucleic Acids Research*, Jan. 1996, pp. 82-85, vol. 24, No. 1.
- Makosky, P. C., et al., "Spectinomycin Resistance at Site 1192 in 16S Ribosomal RNA of *E. coli*: An Analysis of Three Mutants," *Biochimie*, Aug. 1987, pp. 885-889, vol. 69, No. 8.
- McManus, M. C., "Mechanisms of Bacterial Resistance to Antimicrobial Agents," *American Journal of Health-System Pharmacy*, Jun. 1997, pp. 1420-1433, vol. 54, No. 12.

(56)

## References Cited

## OTHER PUBLICATIONS

- Miller, B. T., et al., "Peptide Biotinylation with Amine-Reactive Esters: Differential Side Chain Reactivity," *Peptides*, 1997, pp. 1585-1595, vol. 18, No. 10.
- Miller, J. H., *A Short Course in Bacterial Genetics*, (Miller, J. H. ed.), 1992, pp. 71-80, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Mirza, S. H., et al., "Multi-Drug Resistant Typhoid: A Global Problem," *Journal of Medical Microbiology*, May 1996, pp. 317-319, vol. 44, No. 5.
- Moazed, D., et al., "Interconversion of Active and Inactive 30 S Ribosomal Subunits is Accompanied by a Conformational Change in the Decoding Region of 16 S rRNA," *Journal of Molecular Biology*, Oct. 1986, pp. 483-493, vol. 191, No. 3.
- Morosyuk, S. V., et al., "Structure and Function of the Conserved 690 Hairpin in *Escherichia coli* 16 S Ribosomal RNA. II. NMR Solution Structure," *Journal of Molecular Biology*, Mar. 2001, pp. 197-211, vol. 307, No. 1.
- Morosyuk, S. V., et al., "Structure and Function of the Conserved 690 Hairpin in *Escherichia coli* 16 S Ribosomal RNA. III. Functional Analysis of the 690 Loop," *Journal of Molecular Biology*, Mar. 2001, pp. 213-228, vol. 307, No. 1.
- Morosyuk, S. V., et al., "Structure and Function of the Conserved 690 Hairpin in *Escherichia coli* 16 S Ribosomal RNA: Analysis of the Stem Nucleotides," *Journal of Molecular Biology*, Jun. 2000, pp. 113-126, vol. 300, No. 1.
- Mucha, P., et al., "Anticodon Domain Methylated Nucleosides of Yeast tRNA(Phe) are Significant Recognition Determinants in the Binding of a Phage Display Selected Peptide," *Biochemistry*, Nov. 2001, pp. 14191-14199, vol. 40, No. 47.
- Mucha, P., et al., "Interaction of RNA with Phage Display Selected Peptides Analyzed by Capillary Electrophoresis Mobility Shift Assay," *RNA*, May 2002, pp. 698-704, vol. 8, No. 5.
- Nandi, S., et al., "Gram-Positive Bacteria are a Major Reservoir of Class 1 Antibiotic Resistance Integrons in Poultry Litter," *Proceedings of the National Academy of Sciences of the United States of America*, 2004, pp. 7118-7122, vol. 101, No. 18.
- Nataro, J. P., et al., "Diarrheagenic *Escherichia coli*," *Clinical Microbiology Reviews*, Jan. 1998, pp. 142-201, vol. 11, No. 1.
- Nielsen, D. A., et al., "A Highly Sensitive, Mixed-Phase Assay for Chloramphenicol Acetyltransferase Activity in Transfected Cells," *Analytical Biochemistry*, May 1989, pp. 19-23, vol. 179, No. 1.
- Nikonowicz, E. P., et al., "Preparation of <sup>13</sup>C and <sup>15</sup>N Labelled RNAs for Heteronuclear Multi-Dimensional NMR Studies," *Nucleic Acids Research*, Sep. 1992, pp. 4507-4513, vol. 20, No. 17.
- Ochi, K., "Comparative Ribosomal Protein Sequence Analyses of a Phylogenetically Defined Genus, *Pseudomonas*, and its Relatives," *International Journal of Systematic Bacteriology*, Apr. 1995, pp. 268-273, vol. 45, No. 2.
- O'Connor, M., et al., "Enhancement of Translation by the Epsilon Element is Independent of the Sequence of the 460 Region of 16S rRNA," *Nucleic Acids Research*, Apr. 2001, pp. 1420-1425, vol. 29, No. 7.
- O'Connor, M., et al., "Mutagenesis of the Peptidyltransferase Center of 23S rRNA: The Invariant U2449 is Dispensable," *Nucleic Acids Research*, Feb. 2001, pp. 710-715, vol. 29, No. 3.
- Ogle, J. M., et al., "Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit," *Science*, May 2001, pp. 897-902, vol. 292, No. 5518.
- Oldfield, E. C., 3rd, et al., "The Role of Antibiotics in the Treatment of Infectious Diarrhea," *Gastroenterology Clinics of North America*, Sep. 2001, pp. 817-836, vol. 30, No. 3.
- Orr, J. W., et al., "Protein and Mg(2+)-Induced Conformational Changes in the S15 Binding Site of 16 S Ribosomal RNA," *Journal of Molecular Biology*, Jan. 1998, pp. 453-464, vol. 275, No. 3.
- Papich, M. G., "Antimicrobial Therapy for Gastrointestinal Diseases," *The Veterinary Clinics of North America. Equine Practice*, Dec. 2003, pp. 645-663, vol. 19, No. 3.
- Pelham, H. R., et al., "An Efficient mRNA-Dependent Translation System From Reticulocyte Lysates," *European Journal of Biochemistry*, Aug. 1976, pp. 247-256, vol. 67, No. 1.
- Peske, F., et al., "Conformational Changes of the Small Ribosomal Subunit During Elongation Factor G-Dependent tRNA-mRNA Translocation," *Journal of Molecular Biology*, Nov. 2004, pp. 1183-1194, vol. 343, No. 5.
- Pinkert, C. A., et al., "An Albumin Enhancer Located 10 kb Upstream Functions Along with its Promoter to Direct Efficient, Liver-Specific Expression in Transgenic Mice," *Genes & Development*, May 1987, pp. 268-276, vol. 1, No. 3.
- Pioletti, M., et al., "Crystal Structures of Complexes of the Small Ribosomal Subunit with Tetracycline, Edeine and IF3," *The EMBO Journal*, Apr. 2001, pp. 1829-1839, vol. 20, No. 8.
- Powers, T., et al., "A Functional Pseudoknot in 16S Ribosomal RNA," *The EMBO Journal*, Aug. 1991, pp. 2203-2214, vol. 10, No. 8.
- Prince, J. B., et al., "Covalent Crosslinking of tRNA<sup>Val</sup> to 16S RNA at the Ribosomal P Site: Identification of Crosslinked Residues," *Proceedings of the National Academy of Sciences of the United States of America*, Sep. 1982, pp. 5450-5454, vol. 79, No. 18.
- Queen, C., et al., "Immunoglobulin Gene Transcription is Activated by Downstream Sequence Elements," *Cell*, Jul. 1983, pp. 741-748, vol. 33, No. 3.
- Recht, M. I., et al., "Effect of Mutations in the a Site of 16 S rRNA on Aminoglycoside Antibiotic-Ribosome Interaction," *Journal of Molecular Biology*, Feb. 1999, pp. 33-43, vol. 286, No. 1.
- Rodriguez-Correa, D., et al., "Genetic Evidence Against the 16S Ribosomal RNA Helix 27 Conformational Switch Model," *RNA*, Jan. 2004, pp. 28-33, vol. 10, No. 1.
- Rothman, J. H., et al., "Peptide-Binding Antibiotics: A Solid-Phase Assay for Screening Libraries of Vancomycin Mimics for Selective d-Ala-d-Ala Binding"; *Bioorganic & Medicinal Chemistry Letters*, 1997, pp. 3159-3164, vol. 7.
- Sander, P., et al., "Ribosomal and Non-Ribosomal Resistance to Oxazolidinones: Species-Specific Idiosyncrasy of Ribosomal Alterations," *Molecular Microbiology*, Dec. 2002, pp. 1295-1304, vol. 46, No. 5.
- Schottel, J. L., et al., "Effects of Alterations in the Translation Control Region on Bacterial Gene Expression: Use of Cat Gene Constructs Transcribed from the lac Promoter as a Model System," *Gene*, May 1984, pp. 177-193, vol. 28, No. 2.
- Schultz, L. D., et al., "Expression and Secretion in Yeast of a 400-kDa Envelope Glycoprotein Derived from Epstein-Barr Virus," *Gene*, 1987, pp. 113-123, vol. 54, No. 1.
- Seed, B., "An LFA-3 cDNA Encodes a Phospholipid-Linked Membrane Protein Homologous to its Receptor CD2," *Nature*, 1987, pp. 840-842, vol. 329, No. 6142.
- Serganov, A. A., et al., "The 16S rRNA Binding Site of *Thermophilus thermophilus* Ribosomal Protein S15: Comparison with *Escherichia coli* S15, Minimum Site and Structure," *RNA*, Nov. 1996, pp. 1124-1138, vol. 2, No. 11.
- Sergiev, P. V., et al., "Mutations at Position A960 of *E. coli* 23 S Ribosomal RNA Influence the Structure of 5 S Ribosomal RNA and the Peptidyltransferase Region of 23 S Ribosomal RNA," *Journal of Molecular Biology*, Jun. 2000, pp. 379-389, vol. 299, No. 2.
- Shine, J., et al., "The 3'-Terminal Sequence of *Escherichia coli* 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites," *Proceedings of the National Academy of Sciences of the United States of America*, Apr. 1974, pp. 1342-1346, vol. 71, No. 4.
- Sigmund, C. D., et al., "Antibiotic Resistance Mutations in Ribosomal RNA Genes of *Escherichia coli*," *Methods in Enzymology*, 1988, pp. 673-690, vol. 164.
- Sigmund, C. D., et al., "Antibiotic Resistance Mutations in 16S and 23S Ribosomal RNA Genes of *Escherichia coli*," *Nucleic Acids Research*, Jun. 1984, pp. 4653-4663, vol. 12, No. 11.
- Sigmund, C. D., et al., "Erythromycin Resistance Due to a Mutation in a Ribosomal RNA Operon of *Escherichia coli*," *Proceedings of the National Academy of Sciences of the United States of America*, Sep. 1982, pp. 5602-5606, vol. 79, No. 18.

(56)

**References Cited**

OTHER PUBLICATIONS

Sirinavin, S., et al., "Antibiotics for Treating *Salmonella* Gut Infections," Cochrane Database of Systematic Reviews, 2000, CD001167, No. 2.

Smith, D. B., et al., "Single-Step Purification of Polypeptides Expressed in *Escherichia coli* as Fusions with Glutathione S-Transferase," *Gene*, Jul. 1988, pp. 31-40, vol. 67, No. 1.

Stormo, G. D., et al., "Characterization of Translational Initiation Sites in *E. coli*," *Nucleic Acids Research*, May 1982, pp. 2971-2996, vol. 10, No. 9.

Nucleotide	Description
1-14 & 1499-1536	16S rRNA of <i>Escherichia coli</i> rrnB operon
15-1498	16S rRNA of <i>Pseudomonas aeruginosa</i> rrnC operon
1530-1534	16S MBS (message binding sequence) GGGAU
1537-1976	16S-23S spacer region
1977-4880	23S rRNA of <i>Escherichia coli</i> rrnB operon
4881-4972	23S-5S spacer region
4973-5092	5S rRNA of <i>Escherichia coli</i> rrnB operon
5096-5139	terminator T1 of <i>Escherichia coli</i> rrnB operon
5271-5299	terminator T2 of <i>Escherichia coli</i> rrnB operon
5436-5474	cassette
6462-7319	<i>bla</i> (b-lactamase; ampicillin resistance)
7482-8096	replication origin
8509-8700	<i>rop</i> (Rop protein)
9113-10088	GFP (Green Fluorescent Protein)
10096-10100	GFP RBS (ribosome binding sequence) AUCCC
10117-10157	<i>trpc</i> promoter
10632-10672	<i>trpc</i> promoter
10689-10693	CAT RBS (ribosome binding sequence) AUCCC
10701-11630	<i>cam</i> (chloramphenicol acetyltransferase; CAT)
11670-11747	<i>lacIq</i> promoter
11748-12830	<i>lacIq</i> (lac repressor)
12873-12914	<i>lacUV5</i> promoter

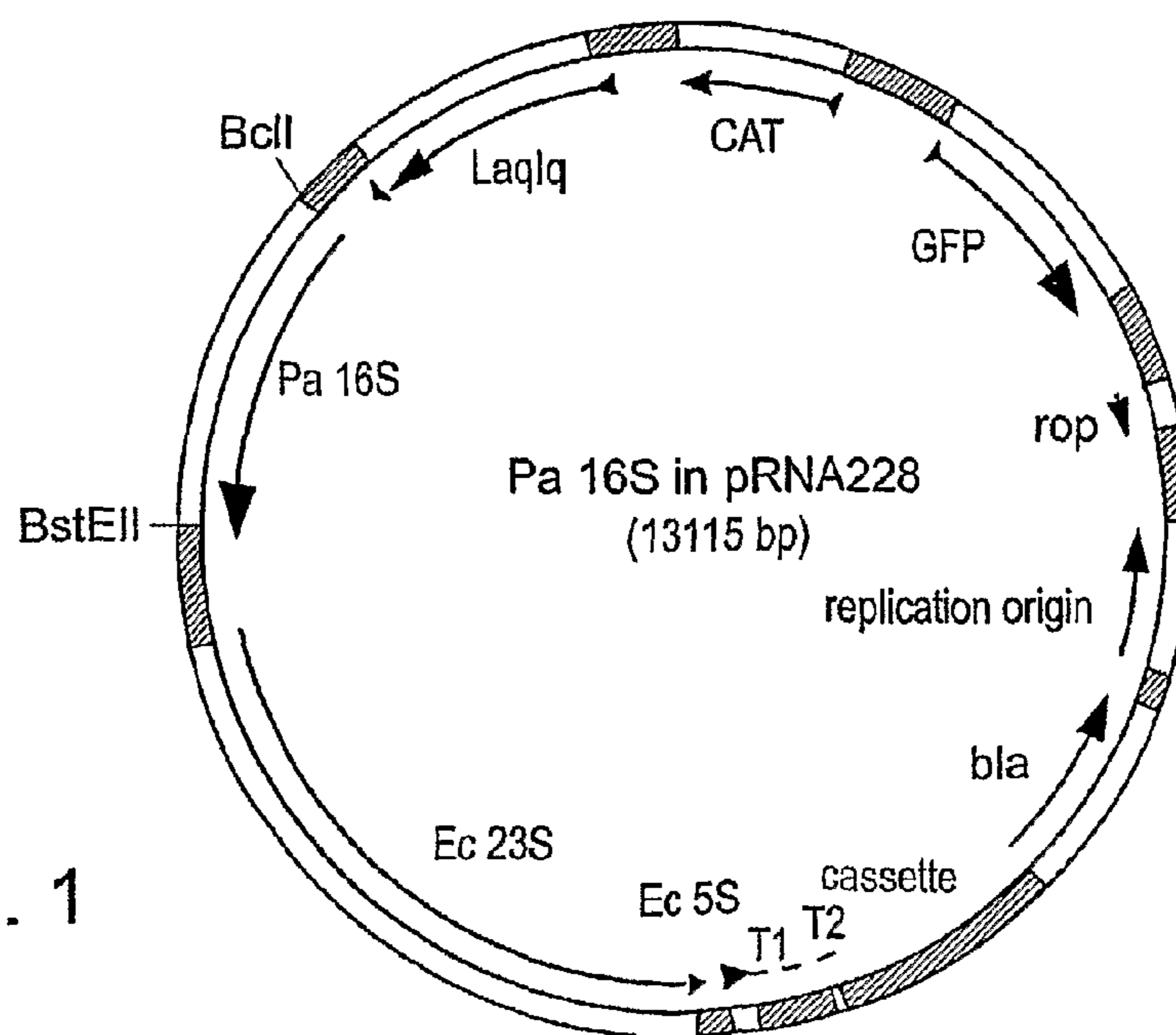


FIG. 1

FIG. 2A  
FIG. 2B

Nucleotide	Description
1-14, 176-188 & 1499-1536	16S rRNA of <i>Escherichia coli</i> rmB operon
15-175 & 189-1498	16S rRNA of <i>Pseudomonas aeruginosa</i> rrmC operon
1530-1534	16S MBS (message binding sequence) GGGAU
1537-1976	16S-23S spacer region
1977-4880	23S rRNA of <i>Escherichia coli</i> rmB operon
4881-4972	23S-5S spacer region
4973-5092	5S rRNA of <i>Escherichia coli</i> rmB operon
5096-5139	terminator T1 of <i>Escherichia coli</i> rmB operon
5271-5299	terminator T2 of <i>Escherichia coli</i> rmB operon
5436-5474	cassette
6462-7319	<i>bla</i> (b-lactamase; ampicillin resistance)
7482-8096	replication origin
8509-8700	<i>rop</i> (Rop protein)
9113-10088	GFP (Green Fluorescent Protein)
10096-10100	GFP RBS (ribosome binding sequence) AUCCC
10117-10157	<i>trpc</i> promoter
10632-10672	<i>trpc</i> promoter
10689-10693	CAT RBS (ribosome binding sequence) AUCCC
10701-11630	<i>cam</i> (chloramphenicol acetyltransferase; CAT)
11670-11747	<i>lacIq</i> promoter
11748-12830	<i>lacIq</i> (lac repressor)
12873-12914	<i>lacUV5</i> promoter

FIG. 2

FIG. 2A



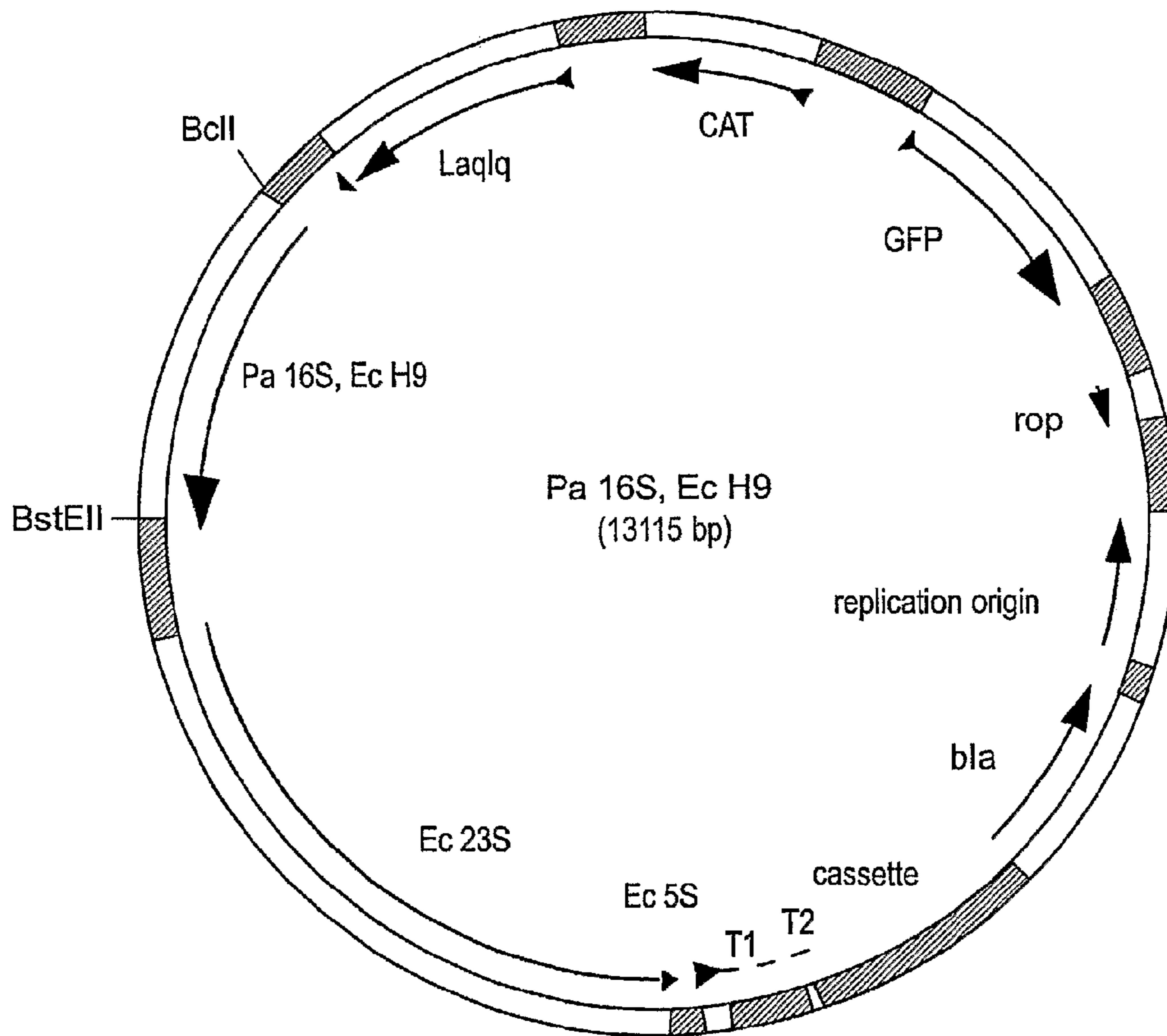


FIG. 2B

Nucleotide	Description
723-1268	replication origin
1862-2677	<i>aph</i> (3')-Ia (kanamycin resistance)
2950-3828	<i>araC</i>
3979-4005	Pc (araC promoter)
4104-4131	pBAD promoter
4163-4206	terminator T1 of <i>Escherichia coli</i> rrnB operon
4338-4366	terminator T2 of <i>Escherichia coli</i> rrnB operon

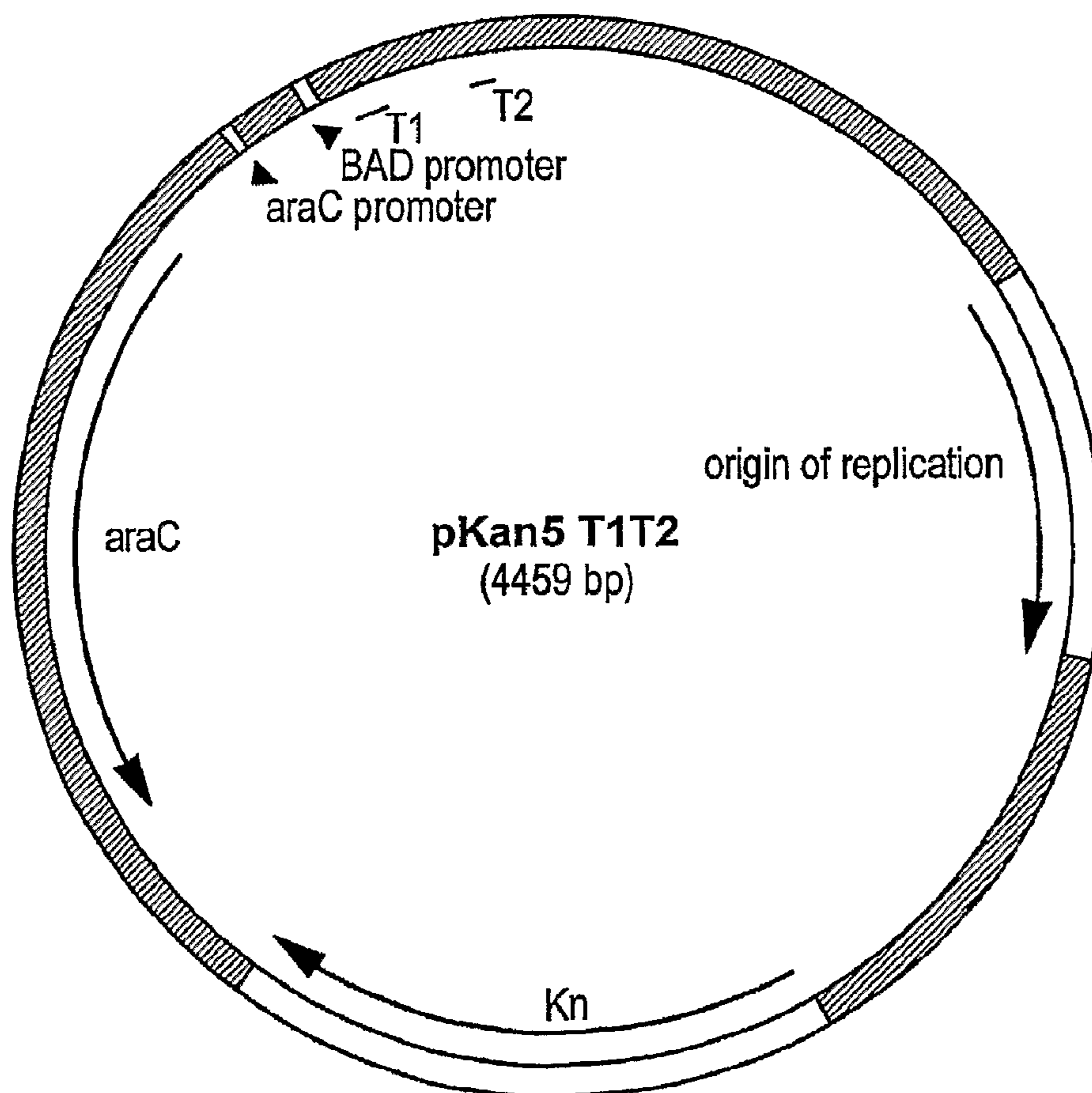


FIG. 3

Nucleotide	Description
723-1268	replication origin
1862-2677	aph (3')-la (kanamycin resistance)
2950-3828	araC
3979-4005	Pc (araC promoter)
4104-4131	pBAD promoter
4168-4172	Shine-Dalgarno sequence (GAGGA)
4181-4453	Pseudomonas aeruginosa S20
4463-4506	terminator T1 of Escherichia coli rrnB operon
4638-4666	terminator T2 of Escherichia coli rrnB operon

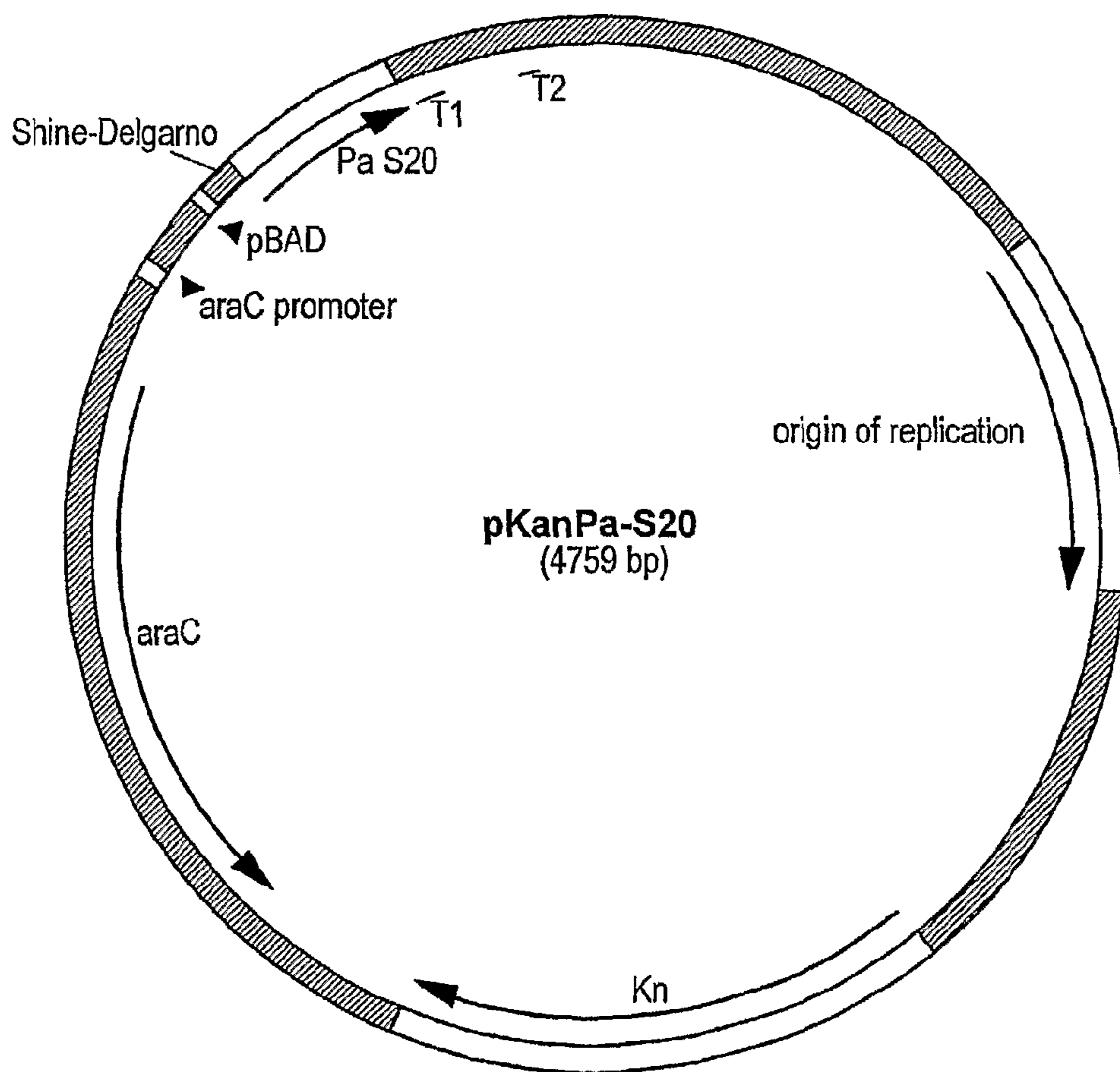


FIG. 4

Figure 5

Samples	GFP %
LB+Amp+Kan+pRNA228 and pKanPa-S20	5.9
LB+Amp+Kan+Pa 16S in pRNA228 and pKanPa-S20	2.38
LB+Amp+Kan+pRNA228 and pKan5-T1T2	5.82
LB+Amp+Kan+Pa 16S in pRNA228 and pKan5-T1T2	1.97
LB+Amp+Kan+IPTG+pRNA228 and pKanPa-S20	83.94
LB+Amp+Kan+IPTG+Pa 16S in pRNA228 and pKanPa-S20	41.57
LB+Amp+Kan+IPTG+pRNA228 and pKan5-T1T2	115.67
LB+Amp+Kan+IPTG+Pa 16S in pRNA228 and pKan5-T1T2	43.05
LB+Amp+Kan+Arabinose+pRNA228 and pKanPa-S20	4.68
LB+Amp+Kan+Arabinose+Pa 16S in pRNA228 and pKanPa-S20	4.07
LB+Amp+Kan+Arabinose+pRNA228 and pKan5-T1T2	4.89
LB+Amp+Kan+Arabinose+Pa 16S in pRNA228 and pKan5-T1T2	1.62
LB+Amp+Kan+IPTG+Arabinose+pRNA228 and pKanPa-S20	100.41
LB+Amp+Kan+IPTG+Arabinose+Pa 16S in pRNA228 and pKanPa-S20	96.85
LB+Amp+Kan+IPTG+Arabinose+pRNA228 and pKan5-T1T2	115.8
LB+Amp+Kan+IPTG+Arabinose+Pa 16S in pRNA228 and pKan5-T1T2	25.04
LB+Amp+IPTG+pRNA228	111.31
LB+Amp+IPTG+Pa 16S in pRNA228	43.55
LB+Amp+IPTG+Arabinose+pRNA228	100
LB+Amp+IPTG+Arabinose+Pa 16S in pRNA228	25.02

Figure 6a

TTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGC  
CGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCA  
GCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAA  
GTTGTCCATATTGGCCACGTTTAAATCAAACACTGGTGAAACTCACCCAGGGATTG  
GCTGAGACGAAAAACATATTCTCAATAAACCCCTTTAGGGAAATAGGCCAGGTTTT  
CACCGTAACACGCCACATCTTGCGAATATATGTGTAGAAACTGCCGGAAATCGTC  
GTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTG  
TAACAAGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATAC  
GGAATTCGGATGAGCATTATCAGGCGGGCAAGAATGTGAATAAAGGCCGGAT  
AAAACCTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAAATATCCAGCTGA  
ACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTT  
TACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATT  
TGCGGAGGGATATGAAAGCGGCCGCTTCCACACATTAACACTAGTTTCGATGATTA  
ATTGTCAACAGCTCGCCGGCGGCACCTCGCTAACGGATTACCACTCCAAGAATT  
GGAGCCAATCGATTCTTGCGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAG  
AACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCA  
GCGTTGGGTCCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTGCTTGAGGACCCG  
GCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGA  
GCGAACGTGAAGCGACTGCTGCTGCAAACGTCTGCGACCTGAGCAACAACATG  
AATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCC  
TGCACCATTATGTTCCGGATCTGGGTACCCGCATTCACAGTTCTCCGCAAGAATC  
GATTGGCTCCAATTCTTGGAGTGGTGAATCCGTTAGCGAGGTGCCGCCGGCGAGC  
TGTTGACAATTAATCATCGAACTAGTTTAAATGTGTGGAAGCGGCCGCTTTCATAT  
CCCTCCGCAAATGGAGAAAAAAATCACTGCTAGCAAAGGAGAAGAACTTTTCAC  
TGGAGTTGTCCCAATTCTTGTGATTAGATGGTGATGTTAATGGGCACAAATTT  
TCTGTCAGTGGAGAGGGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAA  
TTTATTTGCACTACTGGAAAACCTGTTCCATGGCCAAACTTGTCACTACTTT  
CTCTTATGGTGTTCATGCTTTTCCCGTTATCCGGATCATATGAAACGGCATGACT  
TTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAA  
AGATGACGGGAACTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCT  
TGTTAATCGTATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTC  
GGACACAAACTCGAGTACAATACTCACACAATGTATACATCACGGCAGAC  
AAACAAAAGAATGGAATCAAAGCTAACTTCAAATTCGCCACAACATTGAAGAT  
GGATCCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCC  
CTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTGCCCTTTCGAAAGA  
TCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACCTGCTGCTGGG  
ATTACACATGGCATGGATGAGCTCTACAAATAATCTAGTCGTAGCGCCGATGGTA  
GTGTGGGGTCTCCCATGCGAGAGTAGGGAAGTCCAGGCATCAAATAAAACGA  
AAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTGCGGTGAACG  
CTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAAC  
GGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTA  
AGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTTCTACAAACTCTTCCTGT  
CGTCAGTGCAGGTACCGAGCTCGAATCACTGGCCGTCGTTTTACAACGTCGTGA  
CTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTC  
GCCAGGCATCGCAGGATGCTGCTGGCTACCCT

Figure 6b

GTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTT  
TCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAAACGTTCCAGTA  
ACCGGGCATGTTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATC  
GGTATCATTACCCCATGAACAGAAATTCCCCCTTACACGGAGGCATCAAGTGAC  
CAAACAGGAAAAAACCGCCCTTAACATGGCCCCGCTTTATCAGAAGCCAGACATT  
AACGCTTCTGGAGAACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTG  
TGAATCGCTTACGACCACGCTGATGAGCTTACC GCAGCTGCCTCGCGCGTTTC  
GGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCT  
TGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGT  
GTTGGCGGGTGTCCGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTG  
TATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAT  
GCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATAACCGCATCAGGCGCTC  
TTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCG  
GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC  
GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA  
GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAA  
AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAG  
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC  
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTCATAGCTCAC  
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCA  
CGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG  
TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG  
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT  
AACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG  
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG  
TAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT  
CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC  
CACGTTAAGGGATTTTGGTTCATGAGATTATCAAAAAGGATCTTACCTAGATCCT  
TTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG  
TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTAT  
TTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA  
GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCG  
GCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT  
GGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA  
GAGTAAGTAGTTCGCCAGTTAATAGTTTGC GCAACGTTGTTGCCATTGCTGCAGG  
CATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAAC  
GATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTT  
CGGTCCCGATCGTTGTCAGAAGTAAGTTGGCCGCGAGTGTATCACTCATGGTT  
ATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT  
GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGGCACCAG  
TTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGA ACTTTA  
AAAGTGCTCATCATTGGAAAACGTTCTTCCGGGGCGAAAACCTCTCAAGGATCTTAC  
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCCAACTGATCTTCAGC  
ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC  
GCAAAAAGGGGAATAAGGGCGACACGGAA

Figure 6c

ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTT  
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG  
GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATT  
ATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTTCGTCTTCAAG  
AATTCTCATGTTTGACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTATCACA  
GTTAAATTGCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATC  
GTCATCCTCGGCACCGTCACCCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGG  
TACTGCCGGGCTCTTGCGGGATATCGTCCATTCCGACAGCATCGCCAGTCACTA  
TGGCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCG  
GAGCACTGTCCGACCGCTTTGGCCGCCGCCAGTCCCTGCTCGCTTCGCTACTTGG  
AGCCACTATCGACTACGCGATCATGGCGACCACACCCGTCCTGTGGATCCCAGAC  
GAGTTAAGTCACCATACGTTAGTACAGGTTGCCACTCTTTTGGCAGACGCAGACC  
TACGGCTACAATAGCGAAGCGGTCCTGGTATTCATGTTTAAAAATACTGTCGCGA  
TAGCCAAAACGGCACTCTTTGGCAGTTAAGCGCACTTGCTTGCCTGTGCCAGTT  
CAACAGAATCAACATAAGCGCAAACCTCGCTGTAATTTCTACGCCATAAGCACCAA  
TATTCTGGATAGGTGATGAGCCGACACAACCAGGAATTAATGCCAGATTTTCCAG  
ACCAGGCATACCTTCCTGCAAAGTGTATTTTACCAGACGATGCCAGTTTTCTCCG  
GCTCCTACATGTAATAACACGCATCAGGTTTCATCATGAATTTTCGATACCTTTGA  
TCCGGTTGATGATCCCTGCAGGCCCTTAAGGCCATTTAAATGGCGCGCCGATCAA  
TGCCAAATGTGTTCCAGGGTTTTAAGGAGTGGTTCATAGCTGCTTTCTGATGCA  
AAAACGAGGCTAGTTTACCGTATCTGTGGGGGGATGGCTTGTAGATATGACGAC  
AGGAAGAGTTTGTAGAAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTT  
AATTTGATGCCTGGCAGTTTATGGCGGGCGTCCCTGCCCGCCACCCTCCGGGCCGT  
TGCTTCGCAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGT  
TCACCGACAAACAACAGATAAAAACGAAAGGCCAGTCTTTCGACTGAGCCTTTC  
GTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCATGGGGAGACCCACACTAC  
CATCGGCGCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTTCAGGTGGGACCAC  
CGCGCTACTGCCGCCAGGCAAATTCTGTTTTATCAGACCGCTTCTGCGTTCTGATT  
TAATCTGTATCAGGCTGAAAATCTTCTCTCATCCGCCAAAACAGCTTCGGCGTTG  
TAAGGTAAAGCCTCACGGTTCATTAGTACCGGTTAGCTCAACGCATCGCTGCGCT  
TACACACCCGGCCTATCAACGTCGTCGTTCAACGTTCCCTTCAGGACCCTTAAA  
GGGTTCAGGGAGAACTCATCTCGGGGCAAGTTTCGTGCTTAGATGCTTTCAGCACT  
TATCTCTTCCGCATTTAGCTACCGGGCAGTGCCATTGGCATGACAACCCGAACAC  
CAGTGATGCGTCCACTCCGGTCTCTCGTACTAGGAGCAGCCCCCTCAGTTCTC  
CAGCGCCCACGGCAGATAGGGACCGAACTGTCTCACGACGTTCTAAACCCAGCT  
CGCGTACCCTTTAAATGGCGAACAGCCATAACCTTGGGACCTACTTCAGCCCCA  
GGATGTGATGAGCCGACATCGAGGTGCCAAACACCGCCGTCGATATGAACTCTT  
GGGCGGTATCAGCCTGTTATCCCCGGAGTACCTTTTATCCGTTGAGCGATGGCCC  
TTCCATTCAAGAACCCGGATCACTATGACCTGCTTTTCGCACCTGCTCGCGCCGT  
CACGCTCGCAGTCAAGCTGGCTTATGCCATTGCACTAACCTCCTGATGTCCGACC  
AGGATTAGCCAACCTTCGTGCTCCTCCGTTACTCTTTAGGAGGAGACCGCCCCAG  
TCAAACCTACCACAGACTGTCCGCAACCCGGATTACGGGTCAACGTTAGAA  
CATCAAACATTAAGGGTGGTATTTCAAGGTCCGCTCCATGCAGACTGGCGTCCA  
CACTTCAAAGCCTCCACCTATCCTACACATCAAGGCTCAATGTTTCAGTGTCAAG  
CTATAGTAAAGGTTACGGGGTCTTTCGGTCTTGCCG

Figure 6d

CGGGTACACTGCATCTTCACAGCGAGTTCAATTTCACTGAGTCTCGGGTGGAGAC  
AGCCTGGCCATCATTACGCCATTCGTGCAGGTCGGAACCTACCCGACAAGGAATT  
TCGCTACCTTAGGACCGTTATAGTTACGGCCGCCGTTTACCGGGGCTTCGATCAA  
GAGCTTCGCTTGCCTAACCCTCAATTAACCTTCCGGCACCAGGGCAGGCGTCA  
CACCGTATACGTCCACTTTCGTGTTTGCACAGTGCTGTGTTTTTAATAAACAGTTG  
CAGCCAGCTGGTATCTTCGACTGATTTAGCTCCATCCGCGAGGGACCTCACCTA  
CATATCAGCGTGCCTTCTCCCGAAGTTACGGCACCATTTTGCCTAGTTCCTTCACC  
CGAGTTCTCTCAAGCGCCTTGGTATTCTCTACCTGACCACCTGTGTCGGTTTGGGG  
TACGATTTGATGTTACCTGATGCTTAGAGGCTTTTCTGGAAGCAGGGCATTGTG  
GCTTCAGCACCGTAGTGCCTCGTCATCACGCCTCAGCCTTGATTTTCCGGATTTGC  
CTGAAAACCAGCCTACACGCTTAAACCGGGACAACCGTCCCGGGCCAAACATA  
GCCTTCTCCGTCCCCCTTCGCAGTAACACCAAGTACAGGAATATTAACCTGTTT  
CCCATCGACTACGCCTTTCGGCCTCGCCTTAGGGGTGACTCACCTGCCCGGAT  
TAACGTTGGACAGGAACCTTGGTCTTCCGGCGAGCGGGCTTTTACCCGCTTTA  
TCGTTACTTATGTCAGCATTTCGACTTCTGATACCTCCAGCATGCCTCACAGCACA  
CCTTCGCAGGCTTACAGAACGCTCCCCTACCCAACAACGCATAAGCGTCGCTGCC  
GCAGCTTCGGTGCATGGTTTAGCCCCGTTACATCTTCCGCGCAGGGCCGACTCGAC  
CAGTGAGCTATTACGCTTTCTTTAAATGATGGCTGCTTCTAAGCCAACATCCTGG  
CTGTCTGGGCCTTCCACATCGTTTCCCACTTAACCATGACTTTGGGACCTTAGCT  
GGCGGTCTGGGTTGTTTCCCTCTTCACGACGGACGTTAGCACCCGCCGTGTGTCT  
CCCGTGATAACATTCTCCGGTATTCGCAGTTTGCATCGGGTTGGTAAGTCGGGAT  
GACCCCTTGCCGAAACAGTGCTTACCCCGGAGATGAATTCACGAGGCGCTAC  
CTAAATAGCTTTCGGGGAGAACCAGCTATCTCCCGGTTTGATTGGCCTTTCACCC  
CCAGCCACAAGTCATCCGCTAATTTTCAACATTAGTCGGTTCGGTCCCTCCAGTTA  
GTGTTACCCAACCTTCAACCTGCCCATGGCTAGATCACCGGGTTTCGGGTCTATA  
CCCTGCAACTTAACGCCAGTTAAGACTCGGTTTCCCTTCGGCTCCCCTATTCCGGT  
TAACCTTGCTACAGAATATAAGTCGCTGACCCATTATACAAAAGGTACGCAGTCA  
CACGCCTAAGCGTGCTCCCACTGCTTGTACGTACACGGTTTCAGGTTCTTTTTCAC  
TCCCCTCGCCGGGGTCTTTTCGCCTTTCCTCACGGTACTGGTTCCTATCGGTC  
AGTCAGGAGTATTTAGCCTTGGAGGATGGTCCCCCATATTCAGACAGGATACCA  
CGTGTCCCGCCCTACTCATCGAGCTCACAGCATGTGCATTTTTGTGTACGGGGCT  
GTCACCCTGTATCGCGCGCCTTTCAGACGCTTCCACTAACACACACTGATTC  
AGGCTCTGGGCTGCTCCCCGTTTCGCTCGCCGCTACTGGGGGAATCTCGGTTGATT  
TCTTTTCCCTCGGGGTACTTAGATGTTTTCAGTTCCCCCGGTTTCGCCTCATTAACTA  
TGGATTCAGTTAATGATAGTGTGTCGAAACACACTGGGTTTCCCCATTTCGGAAAT  
CGCCGGTTATAACGGTTCATATCACCTTACCGACGCTTATCGCAGATTAGCACGT  
CCTTCATCGCCTCTGACTGCCAGGGCATCCACCGTGTACGCTTAGTCGCTTAACCT  
CACAACCCGAAGATGTTTCTTTCGATTCATCATCGTGTTGCGAAAATTTGAGAGA  
CTCACGAACAACCTCTCGTTGTTTCAGTGTTCATTTTCAGCTTGATCCAGATTTT  
AAAGAGCAAATATCTCAAACATCACCCGAAGATGAGTTTTGAGATATTAAGGTC  
GGCGACTTTCCTCACAAACCAGCAAGTGGCGTCCCCTAGGGGATTTCGAACCCCT  
GTTACCGCCGTGAAAGGGCGGTGTCCTGGGCCTTAGACGAAGGGGACACGAAA  
ATTGCTTATCACGCGTTGCGTGATTTTCGTGTAGGGTGAGCTTTCATTAATAGA  
AAGCGAACGGCCTTATTCTTTCAGCCTCACTCCAACGCGTAAACGCCTTGCTT  
TTCATTTCTATCAGACAATCTGTGTGAGCACTAC



Figure 6e

AAAGTACGCTTCTTTAAGGTAATCCCATGATCCAACCGCAGGTTCCCCTACGGTT  
ACcTTGTTACGACTTCACCCCAGTCATGAATCACTCCGTGGTAACCGTCCCCCTTG  
CGGTTAGACTAGCTACTTCTGGAGCAACCCACTCCCATGGTGTGACGGGCGGTGT  
GTACAAGGCCCGGGAACGTATTCACCGTGACATTCTGATTCACGATTACTAGCGA  
TTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACTACGATCGGTTTTA  
TGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTTTGTACCGACCATTGTAGCAC  
GTGTGTAGCCCTGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCCT  
CCGTTTTGTCACCGGCAGTCTCCTTAGAGTGCCACCCGAGGTGCTGGTAACTAA  
GGACAAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCT  
GACGACAGCCATGCAGCACCTGTGTCTGAGTTCCCGAAGGCACCAATCCATCTCT  
GGAAAGTTCTCAGCATGTCAAGGCCAGGTAAGGTTCTTCGCGTTGCTTCGAATTA  
AACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCATTTGAGTTTTAACCT  
TGCGGCCGTACTCCCCAGGCGGTGACTTATCGCGTTAGCTGCGCCACTAAGATC  
TCAAGGATCCCAACGGCTAGTCGACATCGTTTACGGCGTGGACTACCAGGGTATC  
TAATCCTGTTTTGCTCCCCACGCTTTCGCACCTCAGTGTGAGTATCAGTCCAGGTGG  
TCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTACCGCTACACAGGA  
AATTCCACCACCCTCTACCGTACTCTAGCTCAGTAGTTTTGGATGCAGTTCACG  
GTTGAGCCCCGGGGATTTACATCCAACCTTGCTGAACCACCTACGCGCGCTTTACG  
CCCAGTAATCCGATTAACGCTTGCACCCTTCGTATTACCGCGGCTGCTGGCACG  
AAGTTAGCCGGTGCTTATTCTGTTGGTAACGTCAAACAGCAAGGTATTAACTTA  
CTGCCCTTCCTCCCAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACG  
CGGCATGGCTGGATCAGGCTTTCGCCATTGTCCAATATCCCCACTGCTGCCTCC  
CGTAGGAGTCTGGACCGTGTCTCAGTTCAGTGTGACTGATCATCTCTCAGACC  
AGTTACGGATCGTCGCCTTGGTAGGCCTTTACCCACCAACTAGCTAATCCGACC  
TAGGCTCATCTGATAGCGTGAGGTCCGAAGATCCCCACTTTCTCCCTCAGGACG  
TATGCGGTATTAGCGCCCGTTTTCCGGACGTTATCCCCACTACCAGGCAGATTCC  
TAGGCATTACTCACCCGTCCGCCGCTGAATCCAGGAGCAAGCTCCCTTCATCCGC  
TCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAAAC  
TCTTCAATTTAAAAGTTTGACGCTCAAAGAATTAACCTTCGTAATGAATTACGTG  
TTCACTCTTGAGACTTGGTATTCATTTTTTCGTCTTGCGACGTTAAGAATCCGTATC  
TTCGAGTGCCACACAGATTGTCTGATAAATTGTTAAAGAGCAGTGCCGCTTCGC  
TTTTTCTCAGCGGCCGCTGTGTGAAATTGTTATCCGCTCACAATCCACACATTAT  
ACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACT  
CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGC  
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGGCGTATTGGG  
CGCCAGGGTGGTTTTTTCTTTTACCAGTGAGACGGGCAACAGCTGATTGCCCTTC  
ACCGCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGGCCAGC  
AGGCGAAAATCCTGTTTGATGGTGGTTGACGGCGGGATATAACATGAGCTGTCTT  
CGGTATCGTTCGTATCCCCTACCAGATATCCGCACCAACGCGCAGCCCGGACTC  
GGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCA  
GTGGGAACGATGCCCTCATTAGCATTGTCATGGTTTGTGAAAACCGGACATGG  
CACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATA  
TTTATGCCAGCCAGCCAGACGACGACGCGCCGAGACAGA ACTTAATGGGCCCCGC  
TAACAGCGCGATTTGCTGGTGBCCAATGCGACCAGATGCTCCACGCCAGTCGC  
GTACCGTCTTCATGGGAGAAAATAACT

Figure 6f

GTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCA  
GGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATAATCAGC  
CCACTGACCCGTTGCGCGAGAAGATTGTGCACCGCCGCTTACAGGCTTCGACGC  
CGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGA  
TTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGC  
AACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGA  
ATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGAAAC  
GTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATA  
CTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCT  
CTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCACCATTCGATGGTGTC  
GGATCCTCTAGCCGGACCCACTTGCGGGCCACGATCCGTCCGCGTAAGGCTCATA  
CCGTTAATTATCCCCCCCACGGGAGACCTGAGCAAACCTGCCCTCAGGCATTTGA  
GAAGCACAGGGTCACACTGCTTCGGGTAGTCAATAAACCGGTAAACCAGCAATA  
GACATAAGCGGCTATTTAACGACCCTGCCCTGAACCGACGACCGGGTCGAATTTG  
CTTTCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCACCAG  
GCGTTTAAGGGCACCAATAACTGCCTTAAAAAAA

Figure 7a

TTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGC  
CGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCA  
GCACCTTGTCGCCTTGCCTATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAA  
GTTGTCCATATTGGCCACGTTTAAATCAAACCTGGTGAAACTCACCCAGGGATTG  
GCTGAGACGAAAAACATATTCTCAATAAACCCCTTTAGGGAAATAGGCCAGGTTTT  
CACCGTAACACGCCACATCTTGCGAATATATGTGTAGAAACTGCCGGAAATCGTC  
GTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTG  
TAACAAGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATAC  
GGAATTCCGGATGAGCATTATCAGGCGGGCAAGAATGTGAATAAAGGCCGGAT  
AAAACCTGTGCTTATTTTTCTTTACGGTCTTTAAAAGGCCGTAATATCCAGCTGA  
ACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAATGTTCTT  
TACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATT  
TGCGGAGGGATATGAAAGCGGCCGCTTCCACACATTAAGTTCGATGATTA  
ATTGTCAACAGCTCGCCGGCGGCACCTCGCTAACGGATTCACCACTCCAAGAATT  
GGAGCCAATCGATTCTTGCGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAG  
AACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCA  
GCGTTGGGTCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTGCTTGAGGACCCG  
GCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGA  
GCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAACAACATG  
AATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCC  
TGCACCATTATGTTCCGGATCTGGGTACCCGCATTCACAGTTCTCCGCAAGAATC  
GATTGGCTCCAATTCTTGGAGTGGTGAATCCGTTAGCGAGGTGCCGCCGGCGAGC  
TGTTGACAATTAATCATCGAACTAGTTTAAATGTGTGGAAGCGGCCGCTTTCATAT  
CCCTCCGCAAATGGAGAAAAAATCACTGCTAGCAAAGGAGAAGAAGTTCAC  
TGGAGTTGTCCCAATTCTTGTGTAATTAGATGGTGATGTTAATGGGCACAAATTT  
TCTGTCAAGTGGAGAGGGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAA  
TTTATTTGCACTACTGGAAAACCTGTTCCATGGCCAACACTTGTCACTACTTT  
CTCTTATGGTGTTCATGCTTTTTCCCGTTATCCGGATCATATGAAACGGCATGACT  
TTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAA  
AGATGACGGGAACTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCT  
TGTTAATCGTATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTC  
GGACACAAACTCGAGTACAATACTCACACAATGTATACATCACGGCAGAC  
AAACAAAAGAATGGAATCAAAGCTAACTTCAAATTCGCCACAACATTGAAGAT  
GGATCCGTTCAACTAGCAGACCATTAACAACAAATACTCCAATTGGCGATGGCC  
CTGTCCTTTTACCAGACAACCATTACCTGTGACACAATCTGCCCTTTCGAAAGA  
TCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAAGTGTGCTGGG  
ATTACACATGGCATGGATGAGCTCTACAAATAATCTAGTCGTAGCGCCGATGGTA  
GTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGA  
AAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTGCGGTGAACG  
CTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAAC  
GGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTA  
AGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTCCTGT  
CGTCAGTGCAGGTACCGAGCTCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGA  
CTGGGAAAACCCTGGCGTTACCCAATTAATCGCCTTGACGACACATCCCCCTTC  
GCCAGGCATCGCAGGATGCTGCTGGCTACCCT

Figure 7b

GTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTT  
TCTCTGGTCCC GCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTA  
ACCGGGCATGTTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATC  
GGTATCATTACCCCATGAACAGAAATTCCCCCTTACACGGAGGCATCAAGTGAC  
CAAACAGGAAAAAACCGCCCTTAACATGGCCCCGCTTTATCAGAAGCCAGACATT  
AACGCTTCTGGAGAACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTG  
TGAATCGCTTACGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTC  
GGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCT  
TGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGT  
GTTGGCGGGTGTTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTG  
TATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTA CTGAGAGTGCACCATAT  
GCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCTC  
TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTCGGCTGCGGCGAGCG  
GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC  
GCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAA  
GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAA  
AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG  
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC  
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC  
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCA  
CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG  
TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGG  
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCCT  
AACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG  
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG  
TAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT  
CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC  
CACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCT  
TTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG  
TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTAT  
TTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA  
GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCG  
GCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT  
GGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA  
GAGTAAGTAGTTCCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTGCAGG  
CATCGTGGTGTACGCTCGTTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCTAAC  
GATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT  
CGGTCCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAAGTGTATCACTCATGGTT  
ATGGCAGCACTGCATAATCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT  
GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAG  
TTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTA  
AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTAC  
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC  
ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCC  
GCAAAAAAGGGAATAAGGGCGACACGGAA

Figure 7c

ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTT  
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG  
GGTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATT  
ATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAG  
AATTCTCATGTTTGACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTATCACA  
GTTAAATTGCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAAATGCGCTCATC  
GTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGG  
TACTGCCGGGCCCTTTGCGGGATATCGTCCATTCCGACAGCATCGCCAGTCACTA  
TGGCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCG  
GAGCACTGTCCGACCGCTTTGGCCGCCGCCAGTCTGCTCGCTTCGCTACTTGG  
AGCCACTATCGACTACGCGATCATGGCGACCACACCCGTCCTGTGGATCCCAGAC  
GAGTTAAGTCACCATACGTTAGTACAGGTTGCCACTCTTTTGGCAGACGCAGACC  
TACGGCTACAATAGCGAAGCGGTCTGGTATTCATGTTTAAAAATACTGTGCGGA  
TAGCCAAAACGGCACTCTTTGGCAGTTAAGCGCACTTGCTTGCCTGTGCGCCAGTT  
CAACAGAATCAACATAAGCGCAAACCTCGCTGTAATTTCTACGCCATAAGCACCAA  
TATTCTGGATAGGTGATGAGCCGACACAACCAGGAATTAATGCCAGATTTTCCAG  
ACCAGGCATACCTTTCCTGCAAAGTGTATTTTACCAGACGATGCCAGTTTCTCCG  
GCTCCTACATGTAAATACCACGCATCAGGTTTCATCATGAATTTGATACCTTTGA  
TCCGGTTGATGATCCCTGCAGGCCCTTAAGGCCATTTAAATGGCGCGCCGATCAA  
TGCCAAATGTGTTCCAGGGTTTTAAGGAGTGGTTCATAGCTGCTTTCCTGATGCA  
AAAACGAGGCTAGTTTACCGTATCTGTGGGGGGATGGCTTGTAGATATGACGAC  
AGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTT  
AATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCGCCACCCTCCGGGCCGT  
TGCTTCGCAACGTTCAAATCCGCTCCCGGCGGATTTGTCTACTCAGGAGAGCGT  
TCACCGACAAACAACAGATAAAAACGAAAGGCCAGTCTTTCGACTGAGCCTTTC  
GTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCATGGGGAGACCCACACTAC  
CATCGGCGCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCAC  
CGCGTACTGCCGCCAGGCAAATTTCTGTTTTATCAGACCGCTTCTGCGTTCTGATT  
TAATCTGTATCAGGCTGAAAATCTTCTCTCATCCGCCAAAACAGCTTCGGCGTTG  
TAAGGTTAAGCCTCACGGTTCATTAGTACCGGTTAGCTCAACGCATCGCTGCGCT  
TACACACCCGGCCTATCAACGTCGTCGTCTTCAACGTTCTTTCAGGACCCTTAAA  
GGGTCAGGGAGAACTCATCTCGGGGCAAGTTTCGTGCTTAGATGCTTTCAGCACT  
TATCTCTTCCGCATTTAGCTACCGGGCAGTGCCATTGGCATGACAACCCGAACAC  
CAGTGATGCGTCCACTCCGGTCTCTCGTACTAGGAGCAGCCCCCTCAGTTCTC  
CAGCGCCCACGGCAGATAGGGACCGAACTGTCTCACGACGTTCTAAACCCAGCT  
CGCGTACCCTTTAAATGGCGAACAGCCATAACCCTTGGGACCTACTTCAGCCCCA  
GGATGTGATGAGCCGACATCGAGGTGCCAAACACCGCCGTCGATATGAACTCTT  
GGGCGGTATCAGCCTGTTATCCCCGGAGTACCTTTTATCCGTTGAGCGATGGCCC  
TTCCATTAGAACCCGGATCACTATGACCTGCTTTCGCACCTGCTCGCGCCGT  
CACGCTCGCAGTCAAGCTGGCTTATGCCATTGCACTAACCTCCTGATGTCCGACC  
AGGATTAGCCAACCTTCGTGCTCCTCCGTTACTCTTTAGGAGGAGACCGCCCCAG  
TCAAACCTACCCACCAGACACTGTCCGCAACCCGGATTACGGGTCAACGTTAGAA  
CATCAAACATTAAGGGTGGTATTTCAAGGTCCGCTCCATGCAGACTGGCGTCCA  
CACTTCAAAGCCTCCACCTATCCTACACATCAAGGCTCAATGTTCAAGTGTCAAG  
CTATAGTAAAGGTTACGGGGTCTTTCCGTCTTGCCG

Figure 7d

CGGGTACACTGCATCTTCACAGCGAGTTCAATTTCACTGAGTCTCGGGTGGAGAC  
AGCCTGGCCATCATTACGCCATTCGTGCAGGTCGGAACCTACCCGACAAGGAATT  
TCGCTACCTTAGGACCGTTATAGTTACGGCCGCCGTTTACCGGGGCTTCGATCAA  
GAGCTTCGCTTGCCTAACCCCATCAATTAACCTTCCGGCACCCGGGCAGGCGTCA  
CACCGTATACGTCCACTTTCGTGTTGCACAGTGTGTGTTTTAATAAACAGTTG  
CAGCCAGCTGGTATCTTCGACTGATTTAGCTCCATCCGCGAGGGACCTCACCTA  
CATATCAGCGTGCCTTCTCCCGAAGTTACGGCACCATTTTGCCTAGTTCCTTACC  
CGAGTTCTCTCAAGCGCCTTGGTATTCTCTACCTGACCACCTGTGTCCGTTTGGGG  
TACGATTTGATGTTACCTGATGCTTAGAGGCTTTTCCTGGAAGCAGGGCATTGTT  
GCTTCAGCACCGTAGTGCCTCGTCATCACGCCTCAGCCTTGATTTTCCGGATTGCT  
CTGGAAAACCAGCCTACACGCTTAAACCGGGACAACCGTCGCCCGGCCAACATA  
GCCTTCTCCGTCCCCCTTCGCAGTAACACCAAGTACAGGAATATTAACCTGTTT  
CCCATCGACTACGCCTTTCGGCCTCGCCTTAGGGGTCGACTCACCTGCCCGAT  
TAACGTTGGACAGGAACCCTTGGTCTTCCGGCGAGCGGGCTTTTCACCCGCTTA  
TCGTTACTTATGTCAGCATTTCGCACTTCTGATACCTCCAGCATGCCTCACAGCACA  
CCTTCGCAGGCTTACAGAACGCTCCCTACCCAACAACGCATAAGCGTCGCTGCC  
GCAGCTTCGGTGCATGGTTTAGCCCCGTTACATCTTCCGCGCAGGCCGACTCGAC  
CAGTGAGCTATTACGCTTCTTAAATGATGGCTGCTTCTAAGCCAACATCCTGG  
CTGTCTGGGCCTTCCACATCGTTTCCCACTTAACCATGACTTTGGGACCTTAGCT  
GGCGGTCTGGGTTGTTTCCCTCTTCACGACGGACGTTAGCACCCGCCGTGTGTCT  
CCCGTGATAACATCTCCGGTATTCGCAGTTTGCATCGGGTTGGTAAGTCGGGAT  
GACCCCTTGCCGAAACAGTGTCTACCCCGGAGATGAATTCACGAGGCGCTAC  
CTAAATAGCTTTCGGGGGAGAACCAGCTATCTCCCGGTTTGATTGGCCTTTCACCC  
CCAGCCACAAGTCATCCGCTAATTTTTCAACATTAGTCGGTTCGGTCCCTCCAGTTA  
GTGTTACCCAACCTTCAACCTGCCCATGGCTAGATCACCGGGTTTCGGGTCTATA  
CCCTGCAACTTAACGCCAGTTAAGACTCGGTTTCCCTTCGGCTCCCCTATTCGGT  
TAACCTTGCTACAGAATATAAGTCGCTGACCCATTATACAAAAGGTACGCAGTCA  
CACGCCTAAGCGTGTCTCCCACTGCTTGTACGTACACGGTTTCAGGTTCTTTTTAC  
TCCCCTCGCCGGGGTTCTTTTCGCCTTTCCTCACGGTACTGGTTCACTATCGGTC  
AGTCAGGAGTATTTAGCCTTGGAGGATGGTCCCCCATATTCAGACAGGATACCA  
CGTGTCCCGCCCTACTCATCGAGCTCACAGCATGTGCATTTTTGTGTACGGGGCT  
GTCACCTGTATCGCGCGCCTTTCAGACGCTTCCACTAACACACACTGATTC  
AGGCTCTGGGCTGCTCCCCGTTTCGCTCGCCGCTACTGGGGGAATCTCGGTTGATT  
TCTTTTCCCTCGGGGTACTTAGATGTTTTCAGTTCCCCCGGTTTCGCTCATTAAACCTA  
TGGATTCAGTTAATGATAGTGTGTCGAAACACACTGGGTTTCCCATTTCGGAAAT  
CGCCGGTTATAACGGTTCATATCACCTTACCGACGCTTATCGCAGATTAGCACGT  
CCTTCATCGCCTCTGACTGCCAGGGCATCCACCGTGTACGCTTAGTCGCTTAACCT  
CACAACCCGAAGATGTTTCTTTTCGATTCATCATCGTGTTCGAAAATTTGAGAGA  
CTCACGAACAACCTCTCGTTGTTTCAGTGTTCATTTTCAGCTTGATCCAGATTTTT  
AAAGAGCAAATATCTCAAACATCACCCGAAGATGAGTTTTGAGATATTAAGGTC  
GGCGACTTTCACTCACAAACCAGCAAGTGGCGTCCCCTAGGGGATTCGAACCCCT  
GTTACCGCCGTGAAAGGGCGGTGTCCTGGGCCTTAGACGAAGGGGACACGAAA  
ATTGCTTATCACGCGTTGCGTGATATTTTCGTGTAGGGTGAGCTTTCATTAATAGA  
AAGCGAACGGCCTTATTCTTTCAGCCTCACTCCCAACGCGTAAACGCCTTGCTT  
TTCACTTCTATCAGACAATCTGTGTGAGCACTAC

Figure 7e

AAAGTACGCTTCTTTAAGGTAATCCCATGATCCAACCGCAGGTTCCCCTACGGTT  
ACCTTGTACGACTTCACCCCAGTCATGAATCACTCCGTGGTAACCGTCCCCCTTG  
CGTTAGACTAGCTACTTCTGGAGCAACCCACTCCCATGGTGTGACGGGCGGTGT  
GTACAAGGCCCGGGAACGTATTCACCGTGACATTCTGATTACGATTACTAGCGA  
TTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACTACGATCGGTTTTA  
TGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTTTGTACCGACCATTGTAGCAC  
GTGTGTAGCCCTGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCT  
CCGTTTTGTCACCGGCAGTCTCCTTAGAGTGCCCAACCGAGGTGCTGGTAACTAA  
GGACAAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCT  
GACGACAGCCATGCAGCACCTGTGTCTGAGTTCCCGAAGGCACCAATCCATCTCT  
GGAAAGTTCTCAGCATGTCAAGGCCAGGTAAGGTTCTTCGCGTTGCTTCGAATTA  
AACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCATTTGAGTTTTAACCT  
TGCGGCCGTACTCCCCAGGCGGTGACTTATCGCGTTAGCTGCGCCACTAAGATC  
TCAAGGATCCCAACGGCTAGTCGACATCGTTTACGGCGTGGACTACCAGGGTATC  
TAATCCTGTTTGCTCCCCACGCTTTCGCACCTCAGTGTGAGTATCAGTCCAGGTGG  
TCGCCTTCGCCACTGGTGTTCCTTCTATATCTACGCATTTACCGCTACACAGGA  
AATTCCACCACCCTCTACCGTACTCTAGCTCAGTAGTTTTGGATGCAGTTCCAG  
GTTGAGCCCCGGGATTTCACATCCAACCTTGCTGAACCACCTACGCGCGCTTTACG  
CCCAGTAATTCCGATTAACGCTTGACCCCTTCGTATTACCGCGGCTGCTGGCAGG  
AAGTTAGCCGGTGCTTATTCTGTTGGTAACGTCAAACAGCAAGGTATTAACCTTA  
CTGCCCTTCTCCCAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACG  
CGGCATGGCTGGATCAGGCTTTCGCCCATTTGTCCAATATTCGCCACTGCTGCCTCC  
CGTAGGAGTCTGGACCGTGTCTCAGTTCAGTGTGACTGATCATCCTCTCAGACC  
AGTTACGGATCGTCGCCTTGGTAGGCCTTACCCCACTAGCTAATCCGACC  
TAGGCTCATCTGATAGCGTGAGGTCCGAAGATCCCCACTTTGGTCTTGCGACGT  
TATGCGGTATTAGCGCCCGTTCCGGACGTTATCCCCACTACCAGGCAGATTCC  
TAGGCATTACTACCCGTCCGCCGCTGAATCCAGGAGCAAGCTCCCTTCATCCGC  
TCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAAAC  
TCTTCAATTTAAAAGTTTGACGCTCAAAGAATTAACCTTCGTAATGAATTACGTG  
TTCACTCTTGAGACTTGGTATTCATTTTTCTGCTTTCGCGACGTTAAGAATCCGTATC  
TTCGAGTGCCACACAGATTGTCTGATAAATTGTTAAAGAGCAGTGCCGCTTCGC  
TTTTTCTCAGCGGCCGCTGTGTGAAATTGTTATCCGCTCACAATTCCACACATTAT  
ACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACT  
CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGC  
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCATATTGGG  
CGCCAGGGTGGTTTTTCTTTTACCAGTGAGACGGGCAACAGCTGATTGCCCTTC  
ACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGC  
AGGCGAAAATCCTGTTTGATGGTGGTTGACGGCGGGATATAACATGAGCTGTCTT  
CGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTC  
GGTAATGGCGCGCATTGCGCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCA  
GTGGGAACGATGCCCTCATTAGCATTGTCATGGTTTGTGAAAACCGGACATGG  
CACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATA  
TTTATGCCAGCCAGCCAGACGACGCGCCGAGACAGA ACTTAATGGGCCCCGC  
TAACAGCGCGATTGCTGGTGBCCCAATGCGACCAGATGCTCCACGCCAGTTCGC  
GTACCGTCTTCATGGGAGAAAATAATA

Figure 7f

CTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTG  
CAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATAATCA  
GCCACTGACCCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGAC  
GCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGA  
GATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTG  
GCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGG  
GAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGAA  
ACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCA  
TACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACT  
CTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCACCATTCGATGGTG  
TCGGATCCTCTAGCCGGACCCACTTGCGGCCACGATCCGTCCGCGTAAGGCTCA  
TACCGTTAATTATCCCCCCCACGGGAGACCTGAGCAAACCTGCCCTCAGGCATTT  
GAGAAGCACAGGGTCACACTGCTTCGGGTAGTCAATAAACCGGTAAACCAGCAA  
TAGACATAAGCGGCTATTTAACGACCCTGCCCTGAACCGACGACCGGGTCAATT  
TGCTTTCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCACC  
AGGCGTTTAAGGGCACCAATAACTGCCTTAAAAAAA



Figure 8a

ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTAT  
GCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAC  
GATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGT  
AACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGA  
GCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACCTATTAAC  
TGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCG  
GATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTG  
CTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGG  
GCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC  
AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA  
GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTAAAA  
CTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGAC  
CAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCTTAATAAGA  
TGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAA  
AACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAAC  
CGAGGTAACCTGGCTTGGAGGAGCGCAGTCACCAAACTTGTCTTTTCAGTTTAGC  
CTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAAGTGGCTG  
CTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGATAGTTACC  
GGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTT  
GGAGCGAACTGCCTACCCGGAACCTGAGTGTCAGGCGTGGAATGAGACAAACGCG  
GCCATAACAGCGGAATGACACCGGTAACCGAAAGGCAGGAACAGGAGAGCGC  
ACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGTTTC  
GCCACCACTGATTTGAGCGTCAGATTTCTGTGATGCTTGTGAGGGGGGCGGAGCCT  
ATGGAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCCTGG  
CATCTTCCAGGAAATCTCCGCCCGTTTCGTAAGCCATTTCCGCTCGCCGCAGTCG  
AACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCA  
CATATTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACTTC  
ACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGCTAGCGCT  
GAGGTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATAACCAGGCCTGAATCGCCCC  
ATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGT  
GGACCAGTTGGTGATTTTGAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTGCG  
GGAAGATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTTCGATTTATTCAACAAA  
GCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC  
ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATG  
AGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGG  
ATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTTCGGGCAATCAGGTGC  
GACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACAT  
GGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGG  
CTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGA  
TGCATGGTACTCACTGCGATCCCCGGGAAACAGCATTCCAGGTATTAGAA  
GAATATCCTGATTCAGGTGAAAATATTGCTGATGCGCTGGCAGTGTTCCCTGCGCC  
GGTTGCATTCGATTCCTGTTTGTAAATTGTCCTTTAACAGCGATCGCGTATTTCTG  
CTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTG  
ATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAAC  
TTTTGCCATTCTCACCGGATTCAGTCGTCATC

Figure 8b

ATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG  
TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAACCAGGATCTTGCCATCCTA  
TGGAACCTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAAT  
ATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGA  
GTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTG  
ACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGA  
TCAGATCACGCATCTTCCCGACAACGCAGACCGTTCGGTGGCAAAGCAAAAAGTTC  
AAAATCACTAGTCGACCATGGTACCATCGATGCATAATGTGCCTGTCAAATGGAC  
GAAGCAGGGATTCTGCAAACCCCTATGCTACTCCGTCAAGCCGTCAATTGTCTGAT  
TCGTTACCAATTATGACAACCTTGACGGCTACATCATTCACTTTTTCTTCACAACCG  
GCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAAAT  
AGAGTTGATCGTCAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGG  
TGGTGCTCAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCCTCGCGCCAGCTTAA  
GACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAA  
GCAAACATGCTGTGCGACGCTGGCGATATCAAATGTCTGTCTGCCAGGTGATCG  
CTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGGAGCGACT  
CGTTAATCGCTTCCATGCGCCGCGAGTAACAATTGCTCAAGCAGATTTATCGCCAG  
CAGCTCCGAATAGCGCCCTTCCCTTGCCCGGCGTTAATGATTTGCCCAAACAGG  
TCGCTGAAATGCGGCTGGTGCCTTCATCCGGGCGAAAGAACCCCGTATTGGCA  
AATATTGACGGCCAGTTAAGCCATTTCATGCCAGTAGGCGCGCGGACGAAAGTAA  
ACCCACTGGTGATACCATTGCGGAGCCTCCGGATGACGACCGTAGTGATGAATCT  
CTCCTGGCGGGAAACAGCAAAAATACCCCGGTCCGGCAAACAAATTCCTCGTCCCTG  
ATTTTTACCACCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATT  
CCCAGCGGTTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTT  
AAACCCGCCACCAGATGGGCATTAACGAGTATCCCGGCAGCAGGGGATCATT  
TGCGCTTCAGCCATACTTTTCATACTCCCGCCATTCAGAGAAGAAACCAATTGTC  
CATATTGCATCAGACATTGCCGTCCTGCGTCTTTTACTGGCTCTTCTCGCTAACC  
AAACCGGTAACCCCGCTTATTAAGCATTCTGTAACAAAGCGGGACCAAAGCC  
ATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATT  
GATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATT  
AGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATGCGGCC  
GCATGCAAGCTTCTAGAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTC  
GTTTTATCTGTTGTTTGTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGG  
AGCGGATTTGAACGTTGCGAAGCAACGGCCCCGAGGGTGGCGGGCAGGACGCC  
GCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCT  
TTTTGAATCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGT  
TACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAG

Figure 9a

ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTAT  
GCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTGACAAC  
GATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGT  
AACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAACCAAACGACGA  
GCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACCTATTAAC  
TGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCG  
GATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTCCGGCTGGCTGGTTTATTG  
CTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGG  
GCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC  
AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA  
GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAA  
CTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGAC  
CAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCTTAATAAGA  
TGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAA  
AACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAAC  
CGAGGTAACCTGGCTTGGAGGAGCGCAGTCACCAAACCTTGTCTTTTTCAGTTTAGC  
CTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTGGCTG  
CTGCCAGTGGTGCTTTTTGCATGTCTTTCCGGGTGGACTCAAGACGATAGTTACC  
GGATAAGGCGCAGCGGTCCGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTT  
GGAGCGAACTGCCTACCCGGAACCTGAGTGTCAGGCGTGGAATGAGACAAACGCG  
GCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGC  
ACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTCCGGTTTC  
GCCACCACTGATTTGAGCGTCAGATTTCTGTGATGCTTGTGAGGGGGGCGGAGCCT  
ATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTAAAGTATCTTCTG  
CATCTTCCAGGAAATCTCCGCCCCGTTCTGTAAGCCATTTCCGCTCGCCGAGTCG  
AACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCA  
CATATTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACTTC  
ACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACTCCGCTAGCGCT  
GAGGTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATAACCAGGCCTGAATCGCCCC  
ATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGT  
GGACCAGTTGGTGATTTTGAACCTTTTGCTTTGCCACGGAACGGTCTGCGTTGTCG  
GGAAGATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTTCGATTTATTCAACAAA  
GCCACGTTGTGCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC  
ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATG  
AGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGG  
ATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGC  
GACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACAT  
GGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGG  
CTGACGGAATTTATGCCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGA  
TGCATGGTTACTCACCCTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAA  
GAATATCCTGATTCAGGTGAAAATATTGCTGATGCGCTGGCAGTGTTCTGCGCC  
GGTTGCATTCGATTCCTGTTTGTAAATTGTCTTTTAAACAGCGATCGCGTATTTCTG  
CTCGCTCAGGCGCAATCACGAATGAATAACGGTTTTGGTTGATGCGAGTGATTTTG  
ATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAAC  
TTTTGCCATTCTCACCGGATTCAGTCGTCCTC

Figure 9b

ATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTG  
TATTGATGTTGGACGAGTCGGAATCGCAGACCGATAACCAGGATCTTGCCATCCTA  
TGGAAGTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAAT  
ATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGA  
GTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTG  
ACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGA  
TCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTC  
AAAATCACTAGTCGACCATGGTACCATCGATGCATAATGTGCCTGTCAAATGGAC  
GAAGCAGGGATTCTGCAAACCTATGCTACTCCGTCAAGCCGTCAATTGTCTGAT  
TCGTTACCAATTATGACAACCTTGACGGCTACATCATTCACTTTTTCTTCAACAACCG  
GCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAAAT  
AGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGG  
TGGTGCTCAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAA  
GACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAA  
GCAAACATGCTGTGCGACGCTGGCGATATCAAATTGCTGTCTGCCAGGTGATCG  
CTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGGAGCGACT  
CGTTAATCGCTTCCATGCGCCGCAATAACAATTGCTCAAGCAGATTTATCGCCAG  
CAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTAATGATTTGCCCAAACAGG  
TCGCTGAAATGCGGCTGGTGGCTTTCATCCGGGCGAAAGAACCCCGTATTGGCA  
AATATTGACGGCCAGTTAAGCCATTGATGCCAGTAGGCGCGCGGACGAAAGTAA  
ACCCACTGGTGATAACCAATTCGCGAGCCTCCGGATGACGACCGTAGTGATGAATCT  
CTCCTGGCGGGAACAGCAAAATATCACCCGGTCGGCAAACAATTCTCGTCCCTG  
ATTTTTACCAACCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATT  
CCCAGCGGTTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTT  
AAACCCGCCACCAGATGGGCATTAACGAGTATCCCGGCAGCAGGGGATCATT  
TGCGCTTCAGCCATACTTTTCATACTCCCGCCATTCAGAGAAGAAACCAATTGTC  
CATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACTGGCTCTTCTCGCTAACC  
AAACCGGTAACCCCGCTTATTAAGCATTCTGTAACAAAGCGGGACCAAAGCC  
ATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATT  
GATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATT  
AGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATGCGGCC  
GCCTATTTACCCAGTTTTTTTTGAGGAGCTCGACGATGGCCAACACACCTTCCGC  
CAAAAACGCGCCAAACAGGCTGAGAAGCGTCGACAGCCACAACGCCAGCCTGCG  
CTCCATGGTGCGCACCTACATCAAGAACGTCGTGAAAGCCATCGACGCCAAGGA  
CCTGGAAAAAGCCCAGGCCGCTTACCAGCCGCTGTACCGGTGATCGACCGCAT  
GGCTGACAAAGGCATCATCCACAAGAACAAGGCTGCTCGTCATAAGAGCCGTCT  
GAGCGGCCACATCAAGGCCCTCAGCACCGCTGCCGCTAATCTAGAATAAAACG  
AAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTTCGGTGAAC  
GCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAA  
CGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATT  
AAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGAATTCACTGGCCGTCGTTTTA  
CAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGCAGCAC  
ATCCCCCTTTCGCCAG

Clone	RNA sequences	$\Delta G_{37}^0$ kcal/mol	MIC		CAT	Induction
			$\mu\text{g of Cm/ mL}$	$\mu\text{g of Cm/ mL}$		
Random	5' C A R 1 R 2 R 3 R 4 R 5 C U C G 3' CAT mRNA 3' A U U M 5 M 4 M 3 M 2 M 1 A C U 5' 16S rRNA		-1	+1	-1	+1
pRNA9	5' C A G G A G G C U C G 3' 3' A U U C C U C C A C U 5'	-9.8	500	500	2803 ± 68	2700 ± 196
pRNA6	5' C A G U G U G C U C C G 3' 3' A U U C A C A C A C U 5'	-7.8	100	200	4033 ± 1040	12437 ± 2491
VII30	5' C A U A U C C C U C C G 3' 3' A U U A G G G A C U 5'	-8.4	100	500	6293 ± 706	72206 ± 706
VII43	5' C A A A C A C C C U C G 3' 3' A U U G G A G A C U 5'	-8.1	125	500	5603 ± 1011	47667 ± 891
VII64 VII65	5' C A U A C C C U C C G 3' 3' A U U G G A G A C U 5'	-7.3	100	500	6200 ± 953	37311 ± 3978
VIII29	5' C A U A A U C C U C C G 3' 3' A U U A G G A G A C U 5'	-10.9	125	600	7869 ± 416	91153 ± 4003
VIII46	5' C A A A U A C C U C C G 3' 3' A U U G G A G A C U 5'	-7.7	100	500	6431 ± 816	46840 ± 796
VIII77	5' C A C A U A C C U C C G 3' 3' A U U G G A G A C U 5'	-7.7	150	600	6794 ± 650	44358 ± 4841
VIII93	5' C A C C G A C C U C C G 3' 3' A U U G G A G A C U 5'	-8.5	100	500	5643 ± 897	24888 ± 2388
IX24	5' C A U A U C C C U C C G 3' 3' A U U A G G G A C U 5'	-7.3	100	650	7524 ± 263	91809 ± 4542
IX32	5' C A A C U A C C U C C G 3' 3' A U U G G A G A C U 5'	-7.7	100	500	5783 ± 971	32164 ± 5862
IX67	5' C A U A U A C C U C C G 3' 3' A U U G G A G A C U 5'	-8.1	125	600	6063 ± 787	24581 ± 3009

FIG. 10

Clone	RNA sequences		MIC	
	5' CAUAUCCCU <u>NN</u> NA <u>AA</u> UG3' CAT mRNA	16S rRNA	(µg/mL)	
Mutated positions	3' AUUAGGGUA <u>C</u> UAGG5'		-1	+1
pRNA100	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	100	650
pRNA100 + wt MBS	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	50	50
pRNA122	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	50	600
pRNA122 + wt MBS	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	10	10
pRNA125	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	80	600
pRNA127	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	50	600
pRNA128	5'CAUAUAU <u>CC</u> CU <u>CG</u> AA <u>AG</u> AAUUG3' 3'AUAUA <u>AG</u> GU <u>AG</u> G5'	AAUUG3'	50	600

FIG. 11

FIG. 12A-1
FIG. 12A-2
FIG. 12A-3

FIG. 12a

Clone	Alignment of CAT mRNA and 16S rRNA		MIC ( $\mu\text{g}$ of Cm/mL)		$\Delta G_{37}^0$ (kcal/mol)
	CAT mRNA	16S rRNA	no IPTG	1 mM IPTG	
Random	5'C A <u>R1</u> <u>R2</u> <u>R3</u> <u>R4</u> <u>R5</u> C U C G 3' 3'A U U <u>M5</u> <u>M4</u> <u>M3</u> <u>M2</u> <u>M1</u> A C U 5'				
wild-type	5'C A G G A G G C U C G 3' 3'A U U C C U C C A C U 5'		500	500	-9.8
1	5'C A A U C C C C U C G 3' 3'A U U A G G G A A C U 5'		100	400	-8.3
2	5'C A U A C C U C U C G 3' 3'A U U C C G U A A C U 5'		50	100	-4
3	5'C A C A G U C C U C G 3' 3'A U U A G U A G C A C U 5'		50	100	-1.9

FIG. 12A-1

4	5'C A A A A C C U C G 3' 3'A U U A G U A C U 5'	50	100	-4.1
5	5'C A U A G C C C U C G 3' 3'A U U A U G G G U A C U 5'	50	100	-7.6
6	5'C A U U C U U C C U C G 3' 3'A U U G G A G G A C U 5'	50	100	-7.4
7	5'C A A U U A U C U C G 3' 3'A U U U A A G A C U 5'	50	100	-3.1
8	5'C A C A G A C U C G 3' 3'A U U G A A U A C U 5'	100	100	-3.6
9	5'C A A A G U U C U C G 3' 3'A U U G A G U A C U 5'	100	200	-0.6
10	5'C A U U C A C U C G 3' 3'A U U A A G U A C U 5'	100	400	-7.7
11	5'C A A C U C A C U C G 3' 3'A U U G U G A G A C U 5'	100	200	-7.1
12	5'C A A C C C A C U C G 3' 3'A U U A G G G A C U 5'	50	100	-6
13	5'C A U C C U C U C G 3' 3'A U U G A G A A C U 5'	50	200	-2.2
14	5'C A C A C C A C U C G 3' 3'A U U U G G U A C U 5'	50	100	-4.7

FIG. 12A-2



15	5'C A C C C A C C U C G 3' 3'A U U G G A A C U C U 5'	50	200	-7
16	5'C A U C C C A C U C G 3' 3'A U U G G A A C U C U 5'	50	100	-7.3
17	5'C A A C U C C U C G 3' 3'A U U A U C A U A C U 5'	50	100	-0.8
18	5'C A U A C A U C U C G 3' 3'A U U U G A G A C U 5'	50	100	-2.1
19	5'C A A C U G U G U C G 3' 3'A U U A G A G A C U 5'	50	200	-5.6
20	5'C A A A U A U G U C G 3' 3'A U U U A G A G A C U 5'	200	500	-6.2
21	5'C A U C C U G U C G 3' 3'A U U U G G A G A C U 5'	200	500	-7.3
22	5'C A U A G U A C U C G 3' 3'A U U A G U A C U 5'	100	200	-0.3
23	5'C A A U C C A C U C G 3' 3'A U U A G G U G A C U 5'	200	400	-10.6
24	5'C A C A G A U C U C G 3' 3'A U U U G C G A C U 5'	100	200	-0.2

FIG. 12A-3

FIG. 12B-1
FIG. 12B-2
FIG. 12B-3

FIG. 12b

Clone	Alignment of CAT mRNA and 16S rRNA		MIC ( $\mu\text{g}$ of Cm/mL)		$\Delta G^{\circ}_{37}$ (kcal/mol)
	5'C A R1 R2 R3 R4 R5 C U C G 3'	CAT mRNA 3'A U U M5 M4 M3 M2 M1 A C U 5'	no IPTG	1 mM IPTG	
Random					
25	5'C A U A G C A C U C G 3'	R1 R2 R3 R4 R5 C U C G 3'	200	400	-6.8
	3'A U U A U C G U A C U 5'	M5 M4 M3 M2 M1 A C U 5'			
26	5'C A A C U A A C U C G 3'	R1 R2 R3 R4 R5 C U C G 3'	100	200	-3.4
	3'A U U G U G A U A C U 5'	M5 M4 M3 M2 M1 A C U 5'			
27	5'C A A A U A U C U C G 3'	R1 R2 R3 R4 R5 C U C G 3'	100	400	-5.3
	3'A U U A U G G A A C U 5'	M5 M4 M3 M2 M1 A C U 5'			

FIG. 12B-1

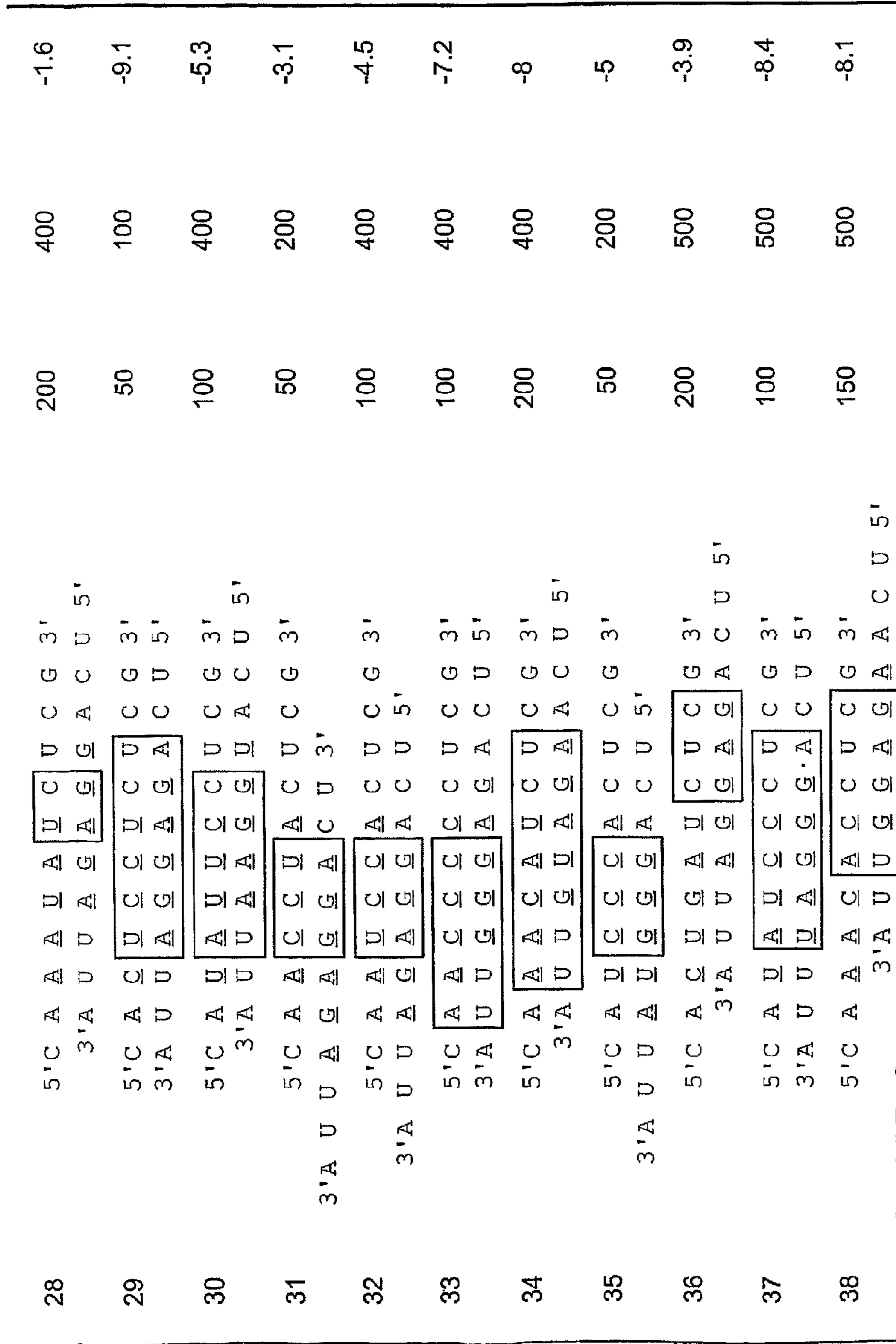


FIG. 12B-2

39	5'C A A C G A A C U C G 3' 3'A U U G U G A G A C U 5'	100	400	-5.7
40	5'C A U C U A U C U C G 3' 3'A U U A G A G G A C U 5'	100	400	-6.2
41	5'C A U C C U C U C G 3' 3'A U U G G A G U A C U 5'	100	500	-7.3
42	5'C A U A U A A C U C G 3' 3'A U U A G A G A A C U 5'	200	500	-3.6
43	5'C A A A U U A C C U C G 3' 3'A U U U G G A G U A C U 5'	100	500	-7.7
44	5'C A C A U U A C C U C G 3' 3'A U U U G G A G U A C U 5'	150	600	-7.7
45	5'C A C C G A C C U C G 3' 3'A U U U G G A G A A C U 5'	100	500	-8.5
46	5'C A U C C C U C C G 3' 3'A U U G G G U A C U 5'	100	700	-7.3
47	5'C A A C U U A C C U C G 3' 3'A U U U G G A G U A C U 5'	100	500	-7.7
48	5'C A U A U A C C U C G 3' 3'A U U U G G A G A A C U 5'	200	600	-8

FIG. 12B-3

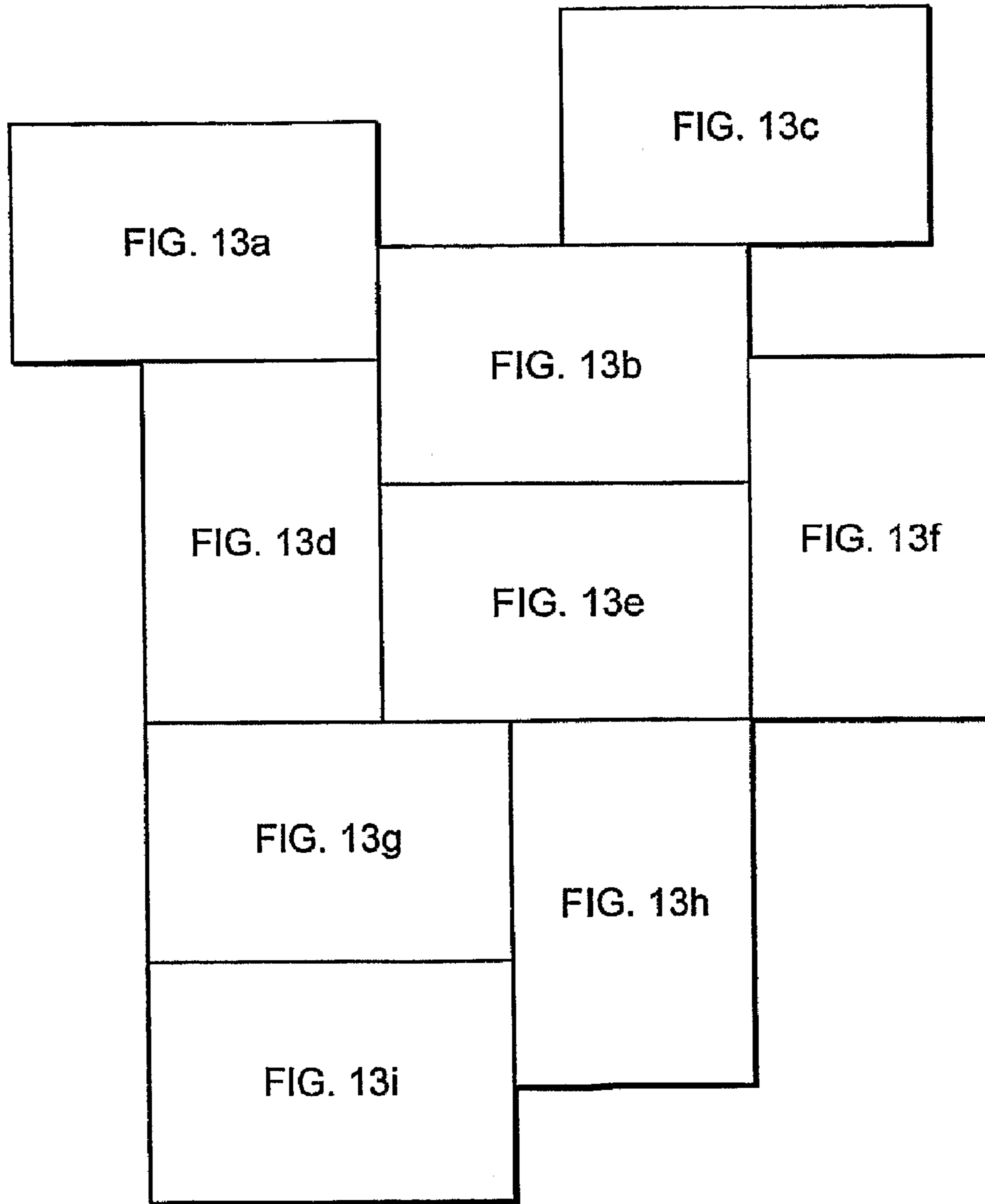


FIG. 13

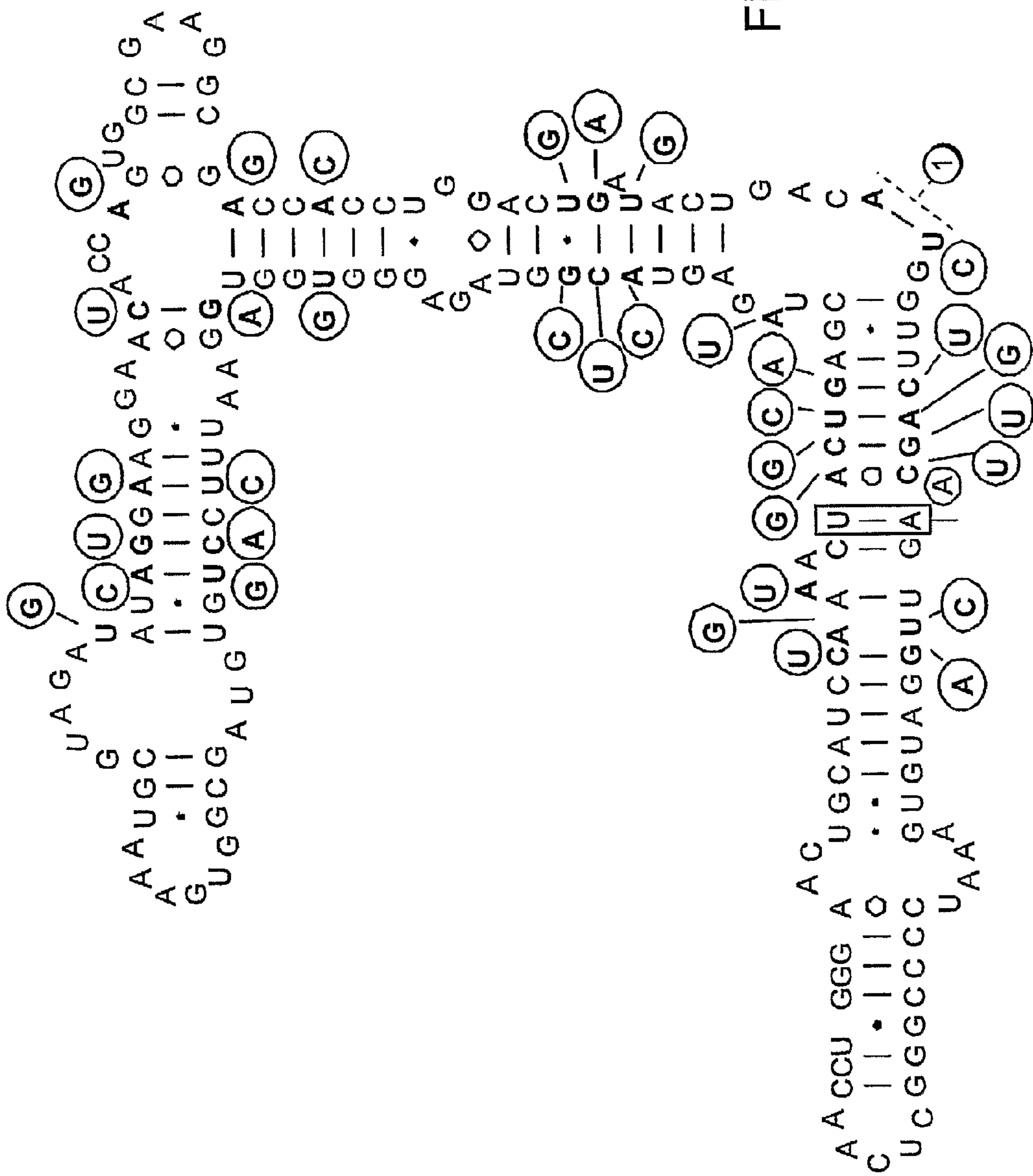


FIG. 13a

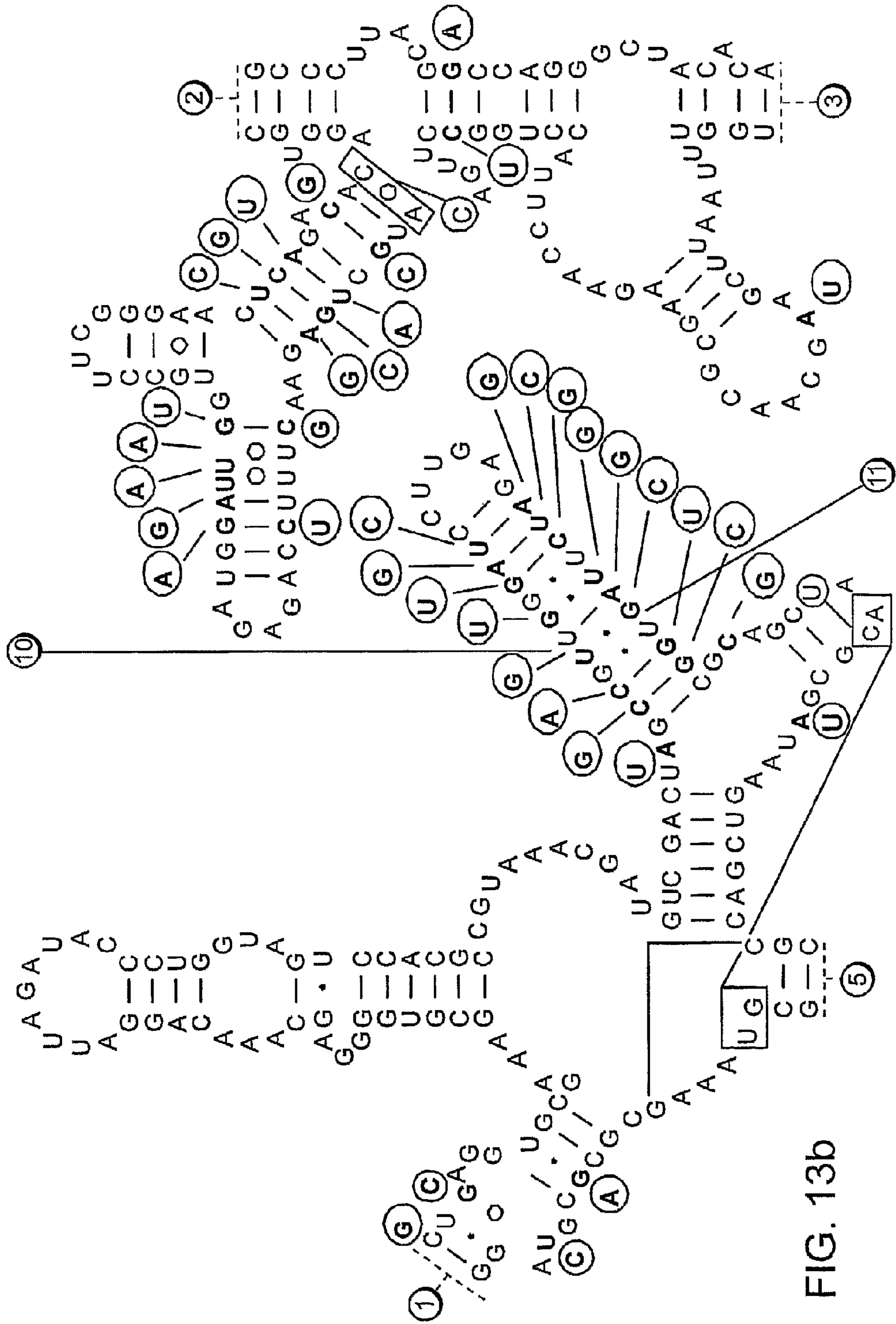


FIG. 13b





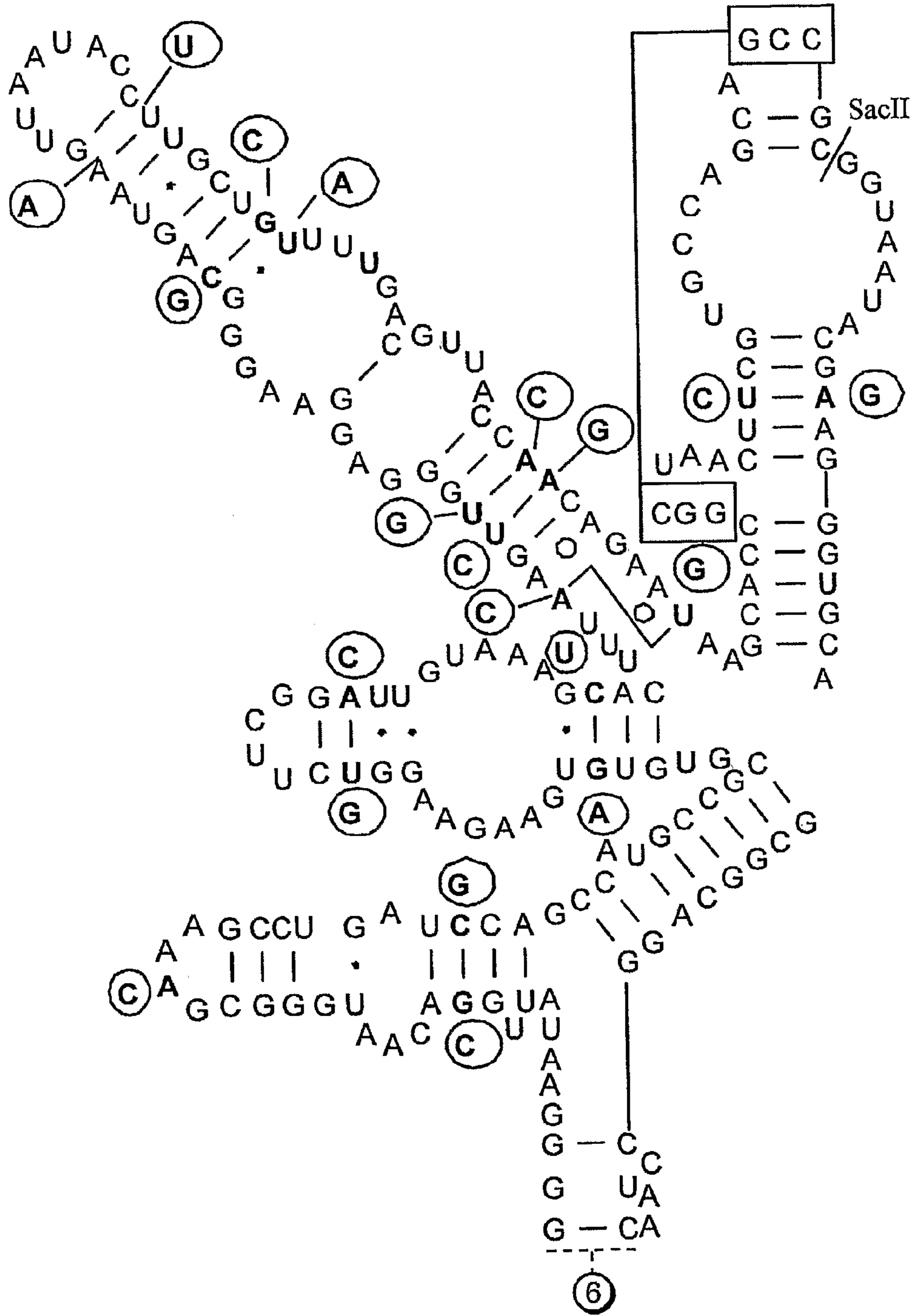


FIG. 13d

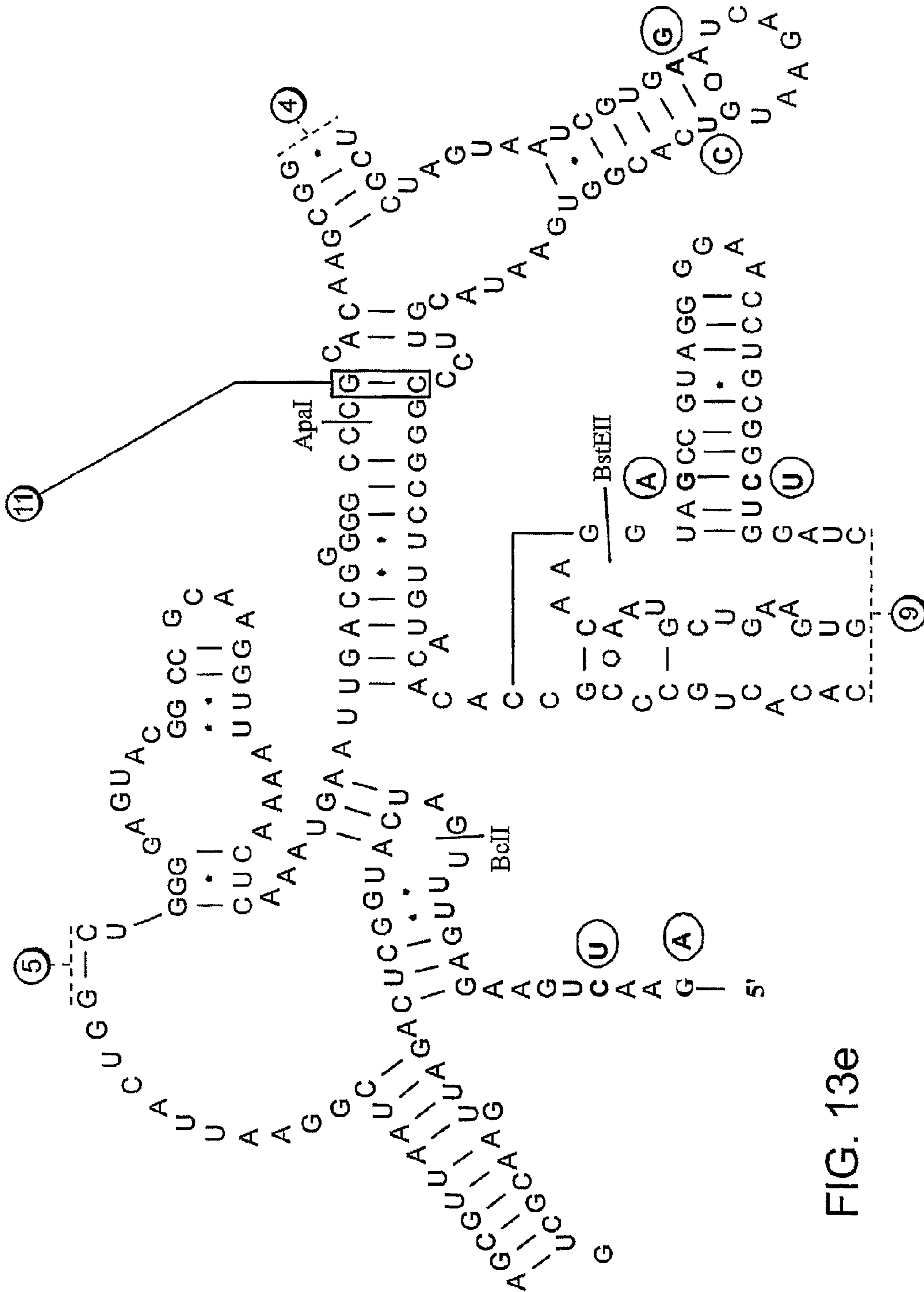


FIG. 13e



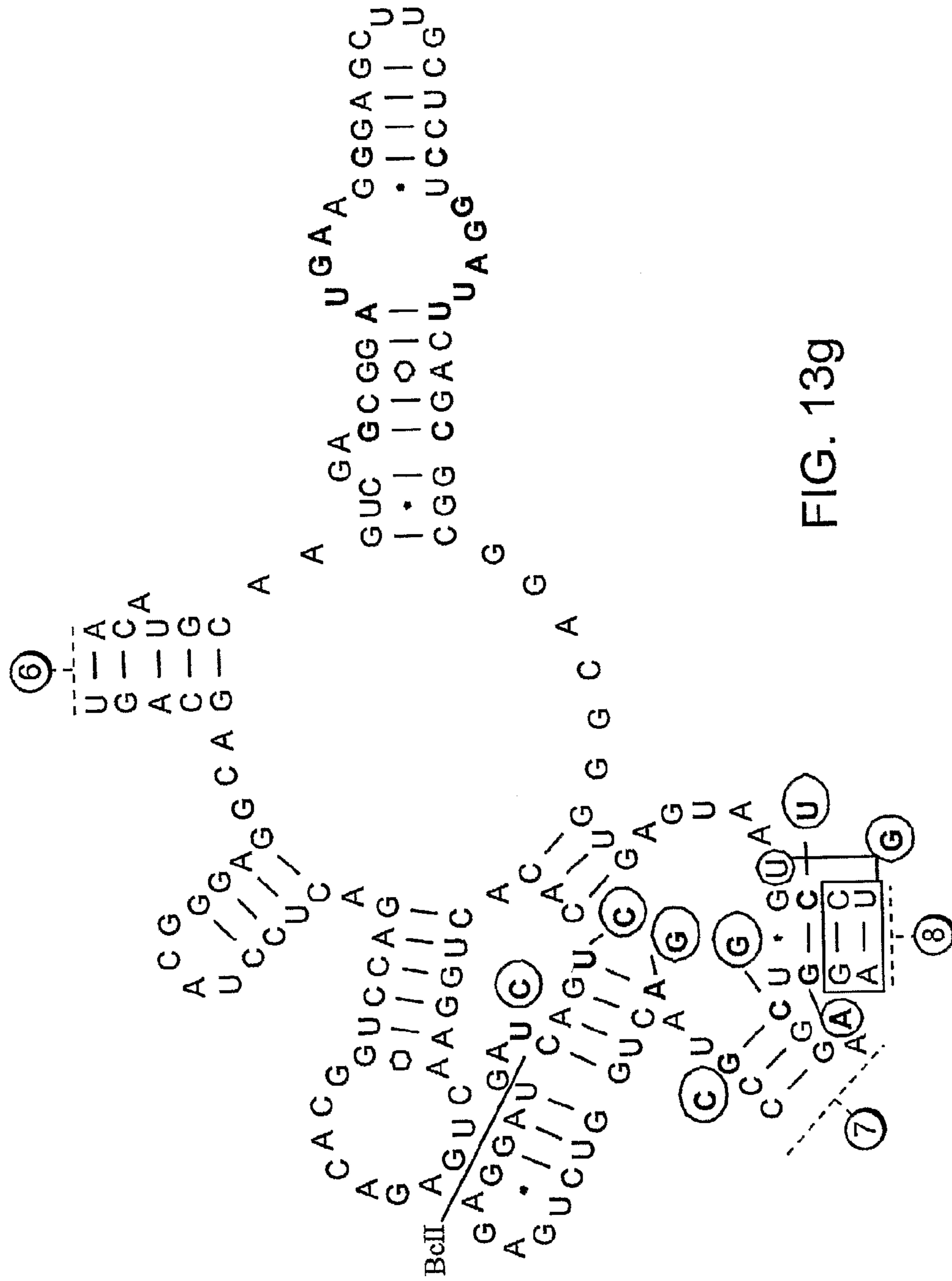


FIG. 13g

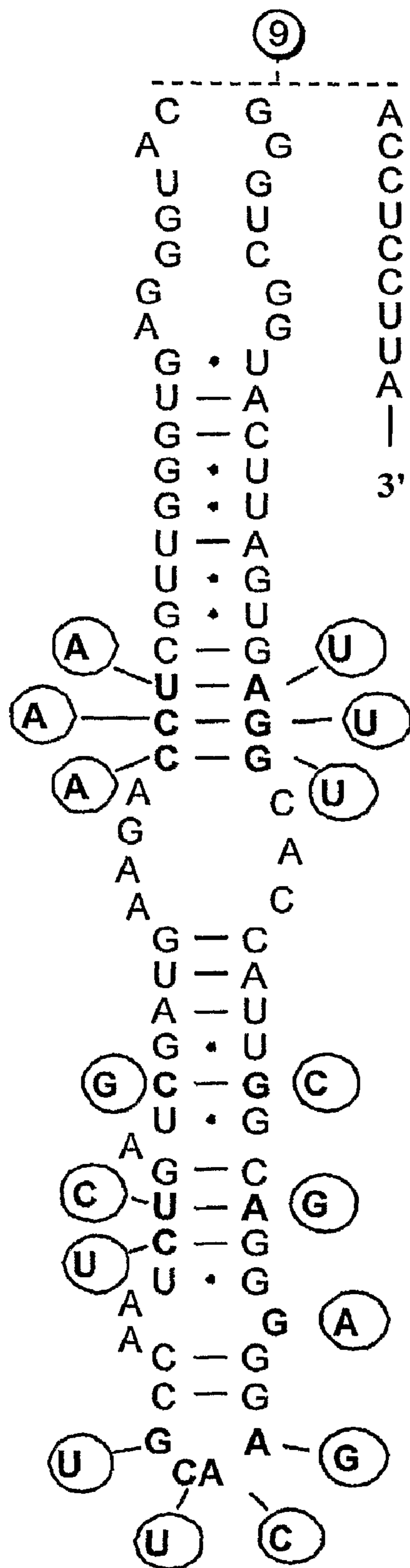


FIG. 13h



Figure 14

<b>E. coli</b>	<b>P. aeruginosa</b>	<b>% GFP</b>
1-14 & 922-1542	15-921	46.2 ± 0.3
1-921 & 1505-1542	922-1504	98.9 ± 0.7
1-528 & 1505-1542	529-1504 basepair positions U513 & A538	121 ± 2.2
1-307 & 1505-1542	308-1504 A289, U307 & U311	108.9 ± 1.9
1-14, 176-188 & 1505-1542	15-175 & 189-1504	96 ± 2.2

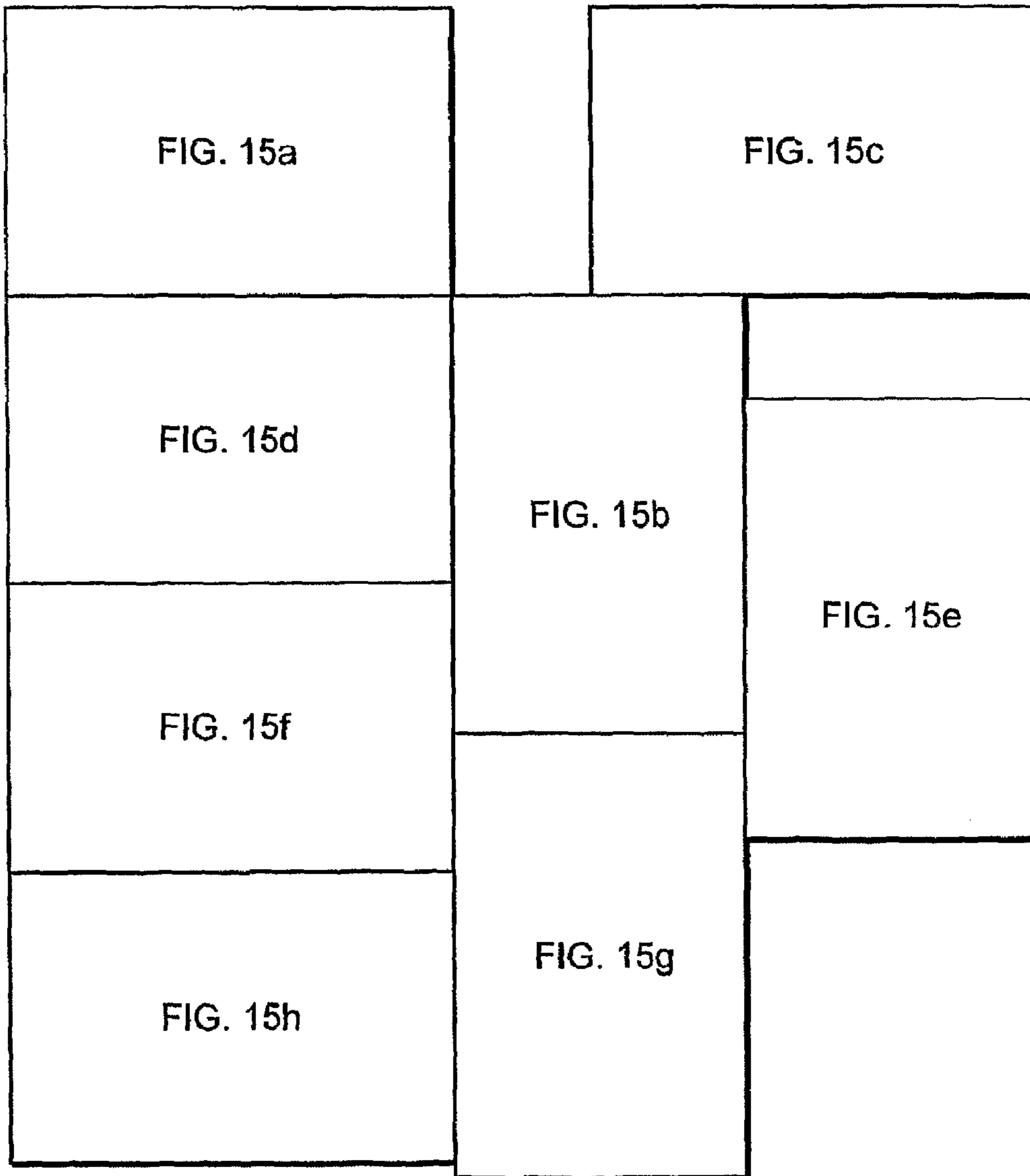


FIG. 15





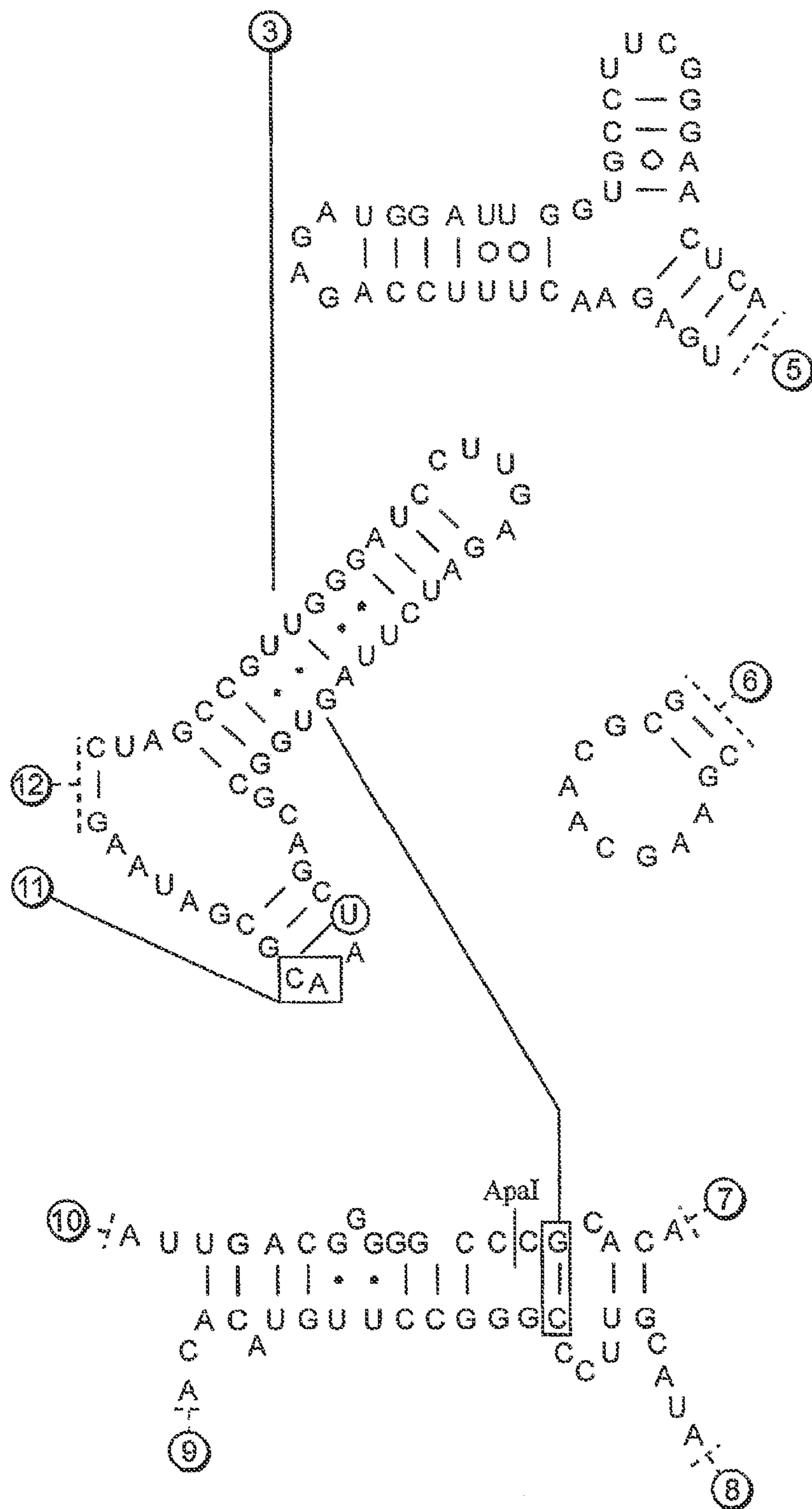


FIG. 15b



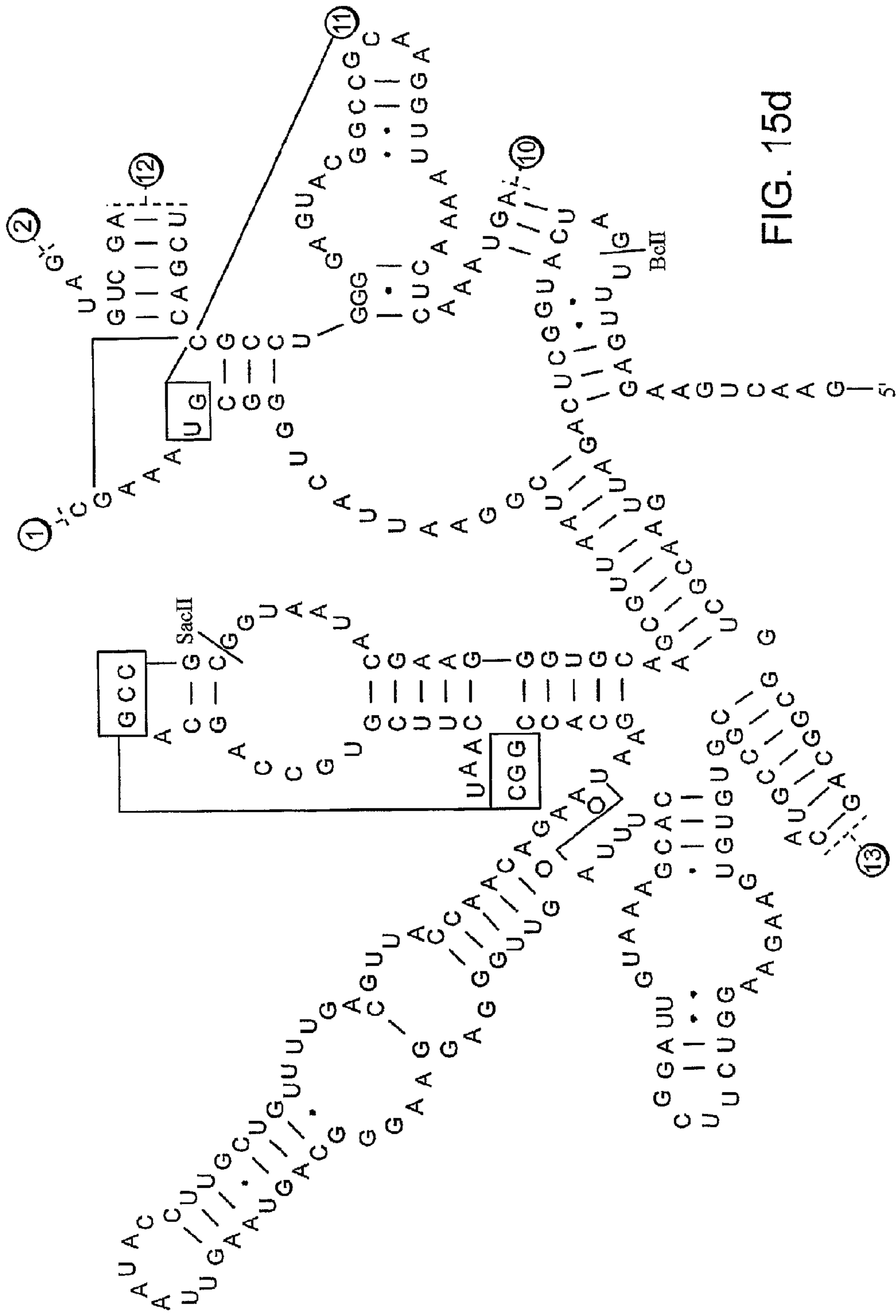


FIG. 15d

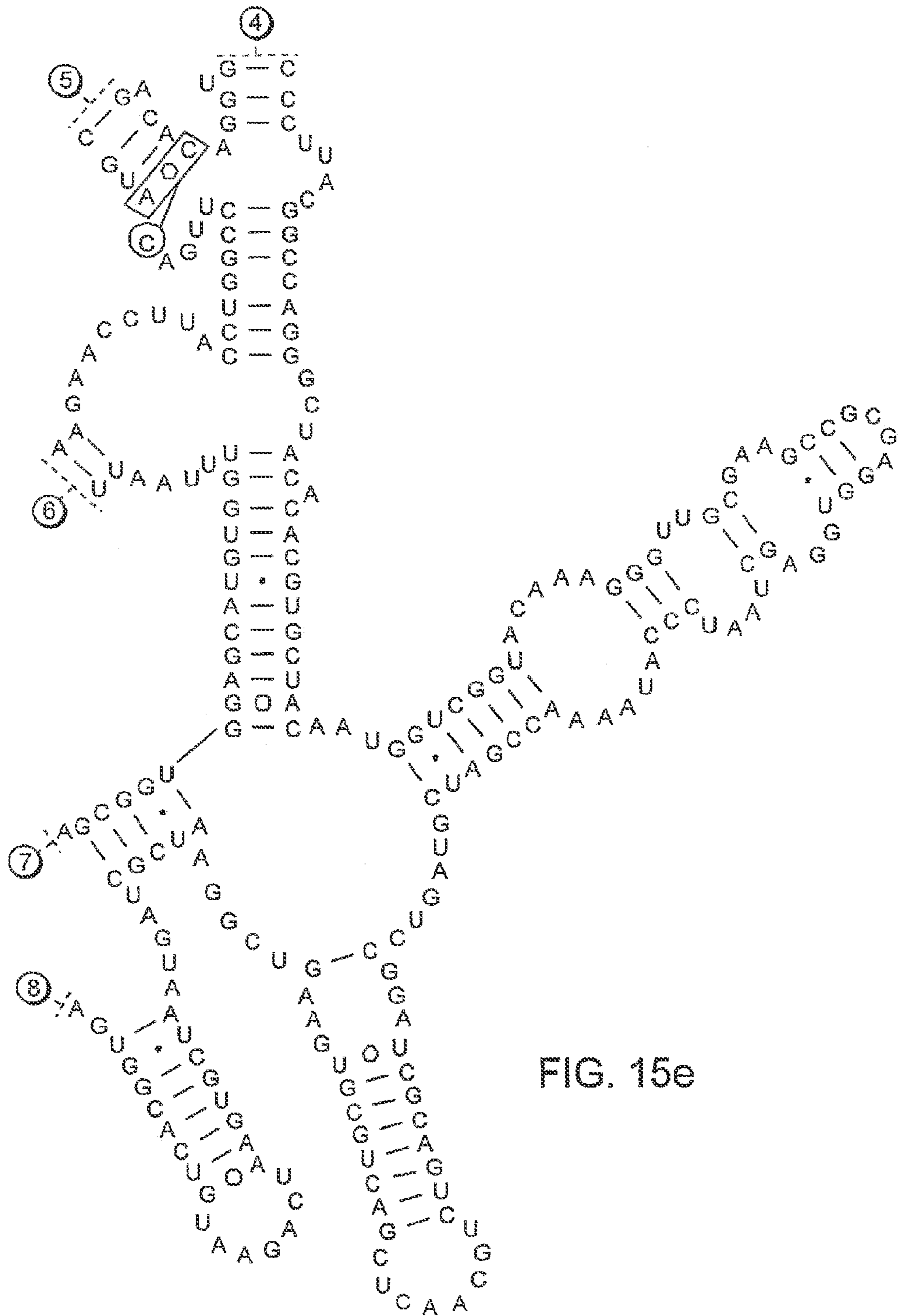


FIG. 15e



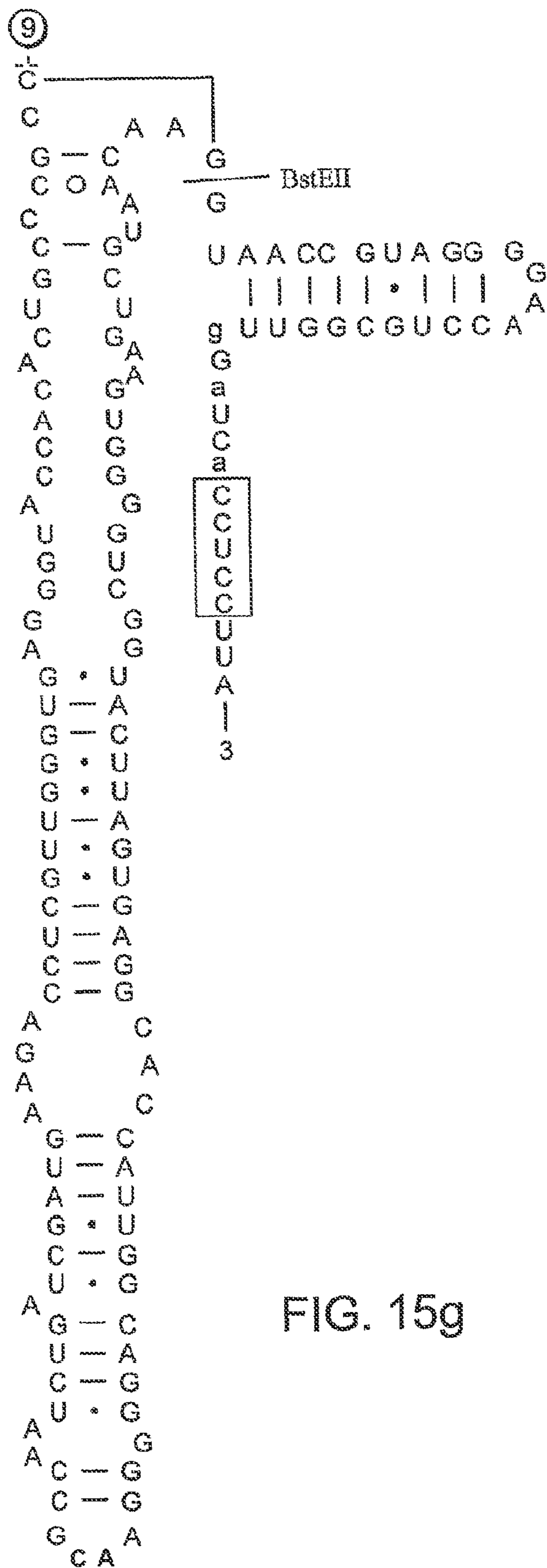
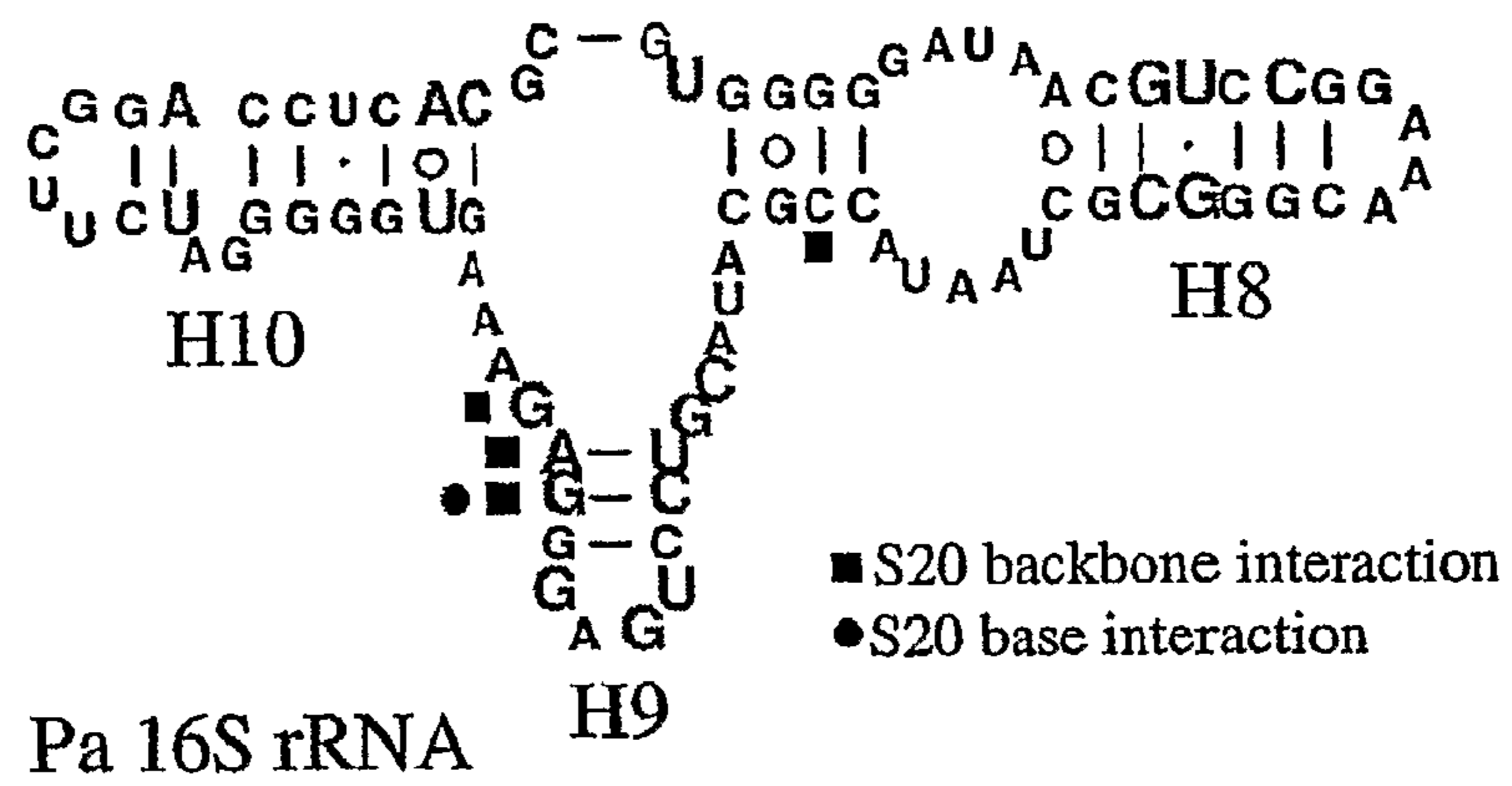


FIG. 15g





Figure 16



Formatted Alignments

	10	20	30
<i>P.a</i>	M A N T P - - - -	S A K K R A K Q A E K R R S	H N A S L R
<i>E. coli</i>	M A N I K - - - -	S A K K R A I Q S E K A R R K	H N A S R R
<i>Thermus</i>	M A Q K K P K R N L S A L K R H R Q S L K R R L	R N K A K K	
	M A N K P K R N L S A K K R A . Q S E K R R	H N A S . R	

	40	50	60
<i>P.a</i>	S M V R T Y I K N V V K A I D A K D L E K I A Q A A F T A A V		
<i>E. coli</i>	S M M R T F I K K V Y A R I E A G D K A A Q K A F N E M Q		
<i>Thermus</i>	S A I K T L S K K A I Q L A Q E G K A E A L K I M R K A E		
	S M . R T . I K K V . . I . A G D . E A Q K A F A		

	70	80	90
<i>P.a</i>	P V I D R M A D K G I I H K N K A A R H K S R L S G H I K A		
<i>E. coli</i>	P I V D R Q A A K G L I H K N K A A R H K A N L T A Q I N K		
<i>Thermus</i>	S L I D K A A K G S T L H K N A A A R R K S R L M R K V R Q		
	P . I D R A K G . I H K N K A A R H K S R L . . . I .		

	100	110	120
<i>P.a</i>	L S T A A A		
<i>E. coli</i>	L A		
<i>Thermus</i>	L L E A A G A P L I G G G L S A		
	L . A A . A P L I G G G L S A		

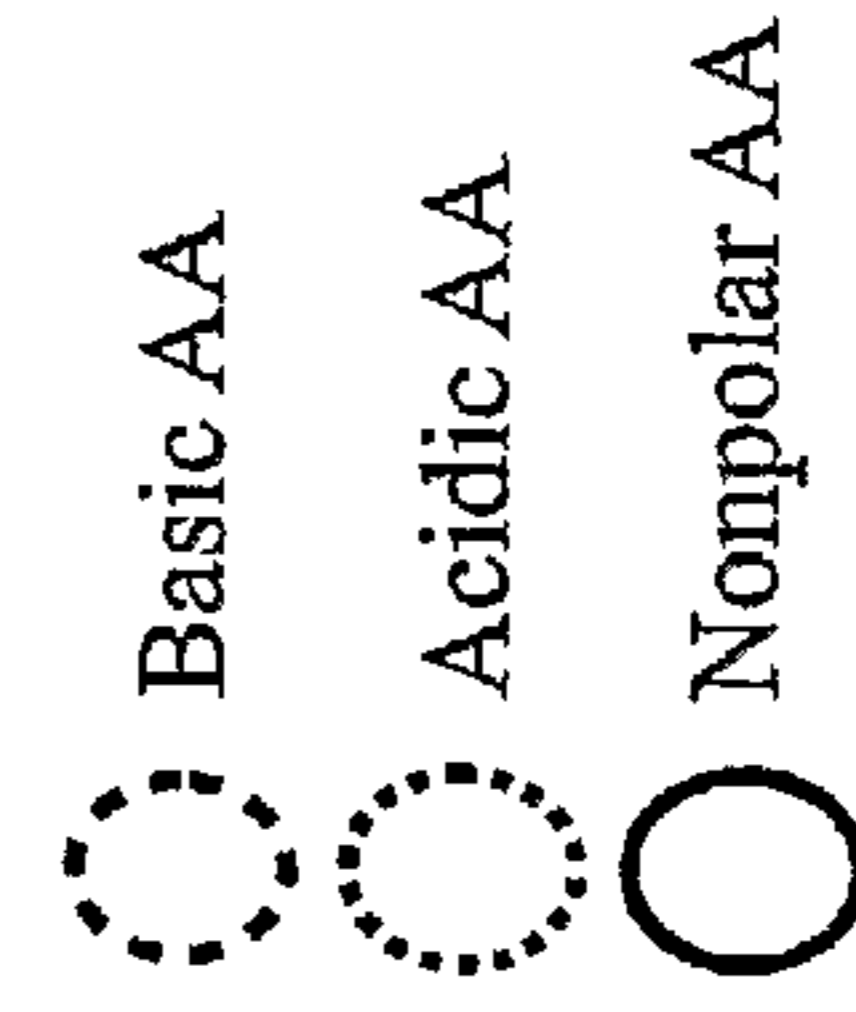


FIG. 17

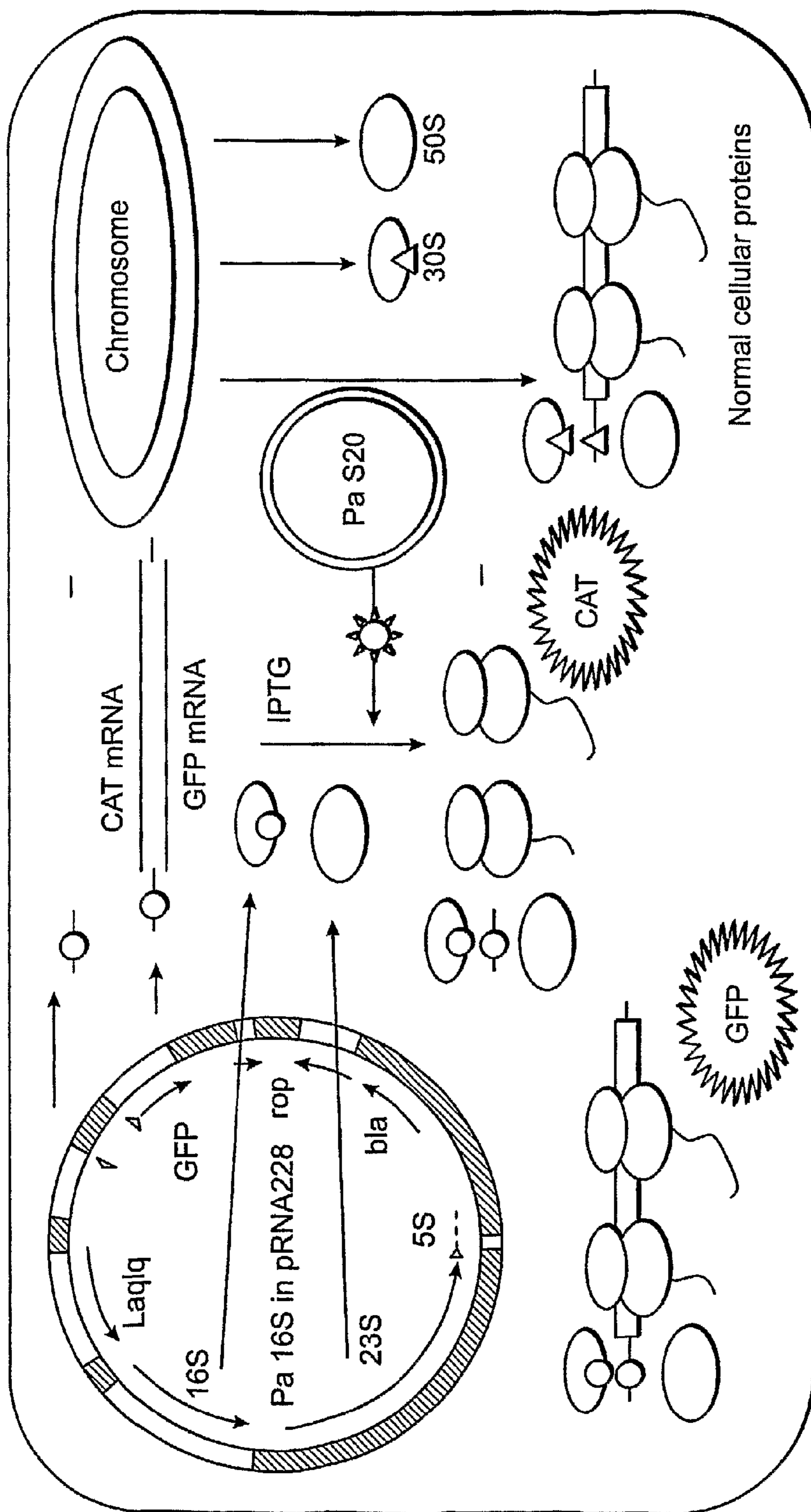


FIG. 18

## 1

**METHODS FOR THE IDENTIFICATION OF  
ANTIBIOTICS THAT ARE NOT  
SUSCEPTIBLE TO ANTIBIOTIC  
RESISTANCE IN *PSEUDOMONAS  
AERUGINOSA***

## RELATED APPLICATIONS

This application is a divisional application of U.S. application Ser. No. 13/654,911, filed Oct. 18, 2012, which is a divisional application of U.S. application Ser. No. 11/914,062, filed Jun. 30, 2008, which is a U.S. national stage application of PCT Patent Application No. PCT/US2006/018187, filed May 11, 2006, which claims the benefit of U.S. Application Ser. No. 60/680,134, filed May 11, 2005, the entire content of each is hereby incorporated herein by reference.

## GOVERNMENT SUPPORT

This invention was made with government support under NIH SBIR Grant Number AI060275-01 awarded by the National Institutes of Health. The government has certain rights in the invention.

## SEQUENCE LISTING

The instant application contains a Sequence Listing which is identical to the Sequence Listing submitted in the parent application Ser. No. 11/914,062, filed on Jun. 30, 2008, via EFS-Web in ASCII format, and is hereby incorporated by reference in its entirety. The ASCII copy is entitled "WSS00401\_SEQUENCE\_LISTING.txt" and is 88012 bytes in size.

## BACKGROUND OF THE INVENTION

Ribosomes are composed of one large and one small subunit containing three or four RNA molecules and over fifty proteins. The part of the ribosome that is directly involved in protein synthesis is the ribosomal RNA (rRNA). The ribosomal proteins are responsible for folding the rRNAs into their correct three-dimensional structures. Ribosomes and the protein synthesis process are very similar in all organisms. One difference between bacteria and other organisms, however, is the way that ribosomes recognize mRNA molecules that are ready to be translated. In bacteria, this process involves a base-pairing interaction between several nucleotides near the beginning of the mRNA and an equal number of nucleotides at the end of the ribosomal RNA molecule in the small subunit. The mRNA sequence is known as the Shine-Dalgarno (SD) sequence and its counterpart on the rRNA is called the Anti-Shine-Dalgarno (ASD) sequence.

There is now extensive biochemical, genetic and phylogenetic evidence indicating that rRNA is directly involved in virtually every aspect of ribosome function (Garrett, R. A., et al. (2000) *The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions*. ASM Press, Washington, D.C.). Genetic and functional analyses of rRNA mutations in *E. coli* and most other organisms have been complicated by the presence of multiple rRNA genes and by the occurrence of dominant lethal rRNA mutations. Because there are seven rRNA operons in *E. coli*, the phenotypic expression of rRNA mutations may be affected by the relative amounts of mutant and wild-type ribosomes in the cell. Thus, detection of mutant phenotypes can be hindered by the presence of

## 2

wild-type ribosomes. A variety of approaches have been designed to circumvent these problems.

One common approach uses cloned copies of a wild-type rRNA operon (Brosius, J., et al. (1981) *Plasmid* 6: 112-118; Sigmund, C. D. et al. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79: 5602-5606). Several groups have used this system to detect phenotypic differences caused by a high level of expression of mutant ribosomes. Recently, a strain of *E. coli* was constructed in which the only supply of ribosomal RNA was plasmid encoded (Asai, T., (1999) *J. Bacteriol.* 181: 3803-3809). This system has been used to study transcriptional regulation of rRNA synthesis, as well as ribosomal RNA function (Voulgaris, J., et al. (1999) *J. Bacteriol.* 181: 4170-4175; Koosha, H., et al. (2000) *RNA*. 6: 1166-1173; Sergiev, P. V., et al. (2000) *J. Mol. Biol.* 299: 379-389; O'Connor, M. et al. (2001) *Nucl. Acids Res.* 29: 1420-1425; O'Connor, M., et al. (2001) *Nucl. Acids Res.* 29: 710-715; Vila-Sanjurjo, A. et al. (2001) *J. Mol. Biol.* 308: 457-463; Morosyuk S. V., et al. (2000) *J. Mol. Biol.* 300 (1):113-126; Morosyuk S. V., et al. (2001) *J. Mol. Biol.* 307 (1):197-210; and Morosyuk S. V., et al. (2001) *J. Mol. Biol.* 307 (1):211-228. Hui et al. showed that mRNA could be directed to a specific subset of plasmid-encoded ribosomes by altering the message binding site (MBS) of the ribosome while at the same time altering the ribosome binding site (RBS) of an mRNA (Hui, A., et al. (1987) *Methods Enzymol.* 153: 432-452).

Although each of the above methods has contributed significantly to the understanding of rRNA function, progress in this field has been hampered both by the complexity of translation and by difficulty in applying standard genetic selection techniques to these systems.

Resistance to antibiotics, a matter of growing concern, is caused partly by antibiotic overuse. According to a study published by the *Journal of the American Medical Association* in 2001, between 1989 to 1999 American adults made some 6.7 million visits a year to the doctor for sore throat. In 73% of those visits, the study found, the patient was treated with antibiotics, though only 5%-17% of sore throats are caused by bacterial infections, the only kind that respond to antibiotics. Macrolide antibiotics in particular are becoming extremely popular for treatment of upper respiratory infections, in part because of their typically short, convenient course of treatment. Research has linked such vast use to a rise in resistant bacteria and the recent development of multiple drug resistance has underscored the need for antibiotics which are highly specific and refractory to the development of drug resistance.

Microorganisms can be resistant to antibiotics by four mechanisms. First, resistance can occur by reducing the amount of antibiotic that accumulates in the cell. Cells can accomplish this by either reducing the uptake of the antibiotic into the cell or by pumping the antibiotic out of the cell. Uptake mediated resistance often occurs, because a particular organism does not have the antibiotic transport protein on the cell surface or occasionally when the constituents of the membrane are mutated in a way that interferes with transport of the antibiotic into a cell. Uptake mediated resistance is only possible in instances where the drug gains entry through a nonessential transport molecule. Efflux mechanisms of antibiotic resistance occur via transporter proteins. These can be highly specific transporters that transport a particular antibiotic, such as tetracycline, out of the cell or they can be more general transporters that transport groups of molecules with similar characteristics out of the cell. The most notorious example of a nonspecific transporter is the multi drug resistance transporter (MDR).

Inactivating the antibiotic is another mechanism by which microorganisms can become resistant to antibiotics. Antibiotic inactivation is accomplished when an enzyme in the cell chemically alters the antibiotic so that it no longer binds to its intended target. These enzymes are usually very specific and have evolved over millions of years, along with the antibiotics that they inactivate. Examples of antibiotics that are enzymatically inactivated are penicillin, chloramphenicol, and kanamycin.

Resistance can also occur by modifying or overproducing the target site. The target molecule of the antibiotic is either mutated or chemically modified so that it no longer binds the antibiotic. This is possible only if modification of the target does not interfere with normal cellular functions. Target site overproduction is less common but can also produce cells that are resistant to antibiotics.

Lastly, target bypass is a mechanism by which microorganisms can become resistant to antibiotics. In bypass mechanisms, two metabolic pathways or targets exist in the cell and one is not sensitive to the antibiotic. Treatment with the antibiotic selects cells with more reliance on the second, antibiotic-resistant pathway.

Among these mechanisms, the greatest concern for new antibiotic development is target site modification. Enzymatic inactivation and specific transport mechanisms require the existence of a substrate specific enzyme to inactivate or transport the antibiotic out of the cell. Enzymes have evolved over millions of years in response to naturally occurring antibiotics. Since microorganisms cannot spontaneously generate new enzymes, these mechanisms are unlikely to pose a significant threat to the development of new synthetic antibiotics. Target bypass only occurs in cells where redundant metabolic pathways exist. As understanding of the MDR transporters increases, it is increasingly possible to develop drugs that are not transported out of the cell by them. Thus, target site modification poses the greatest risk for the development of antibiotic resistance for new classes of antibiotic and this is particularly true for those antibiotics that target ribosomes. The only new class of antibiotics in thirty-five years, the oxazolidinones, is a recent example of an antibiotic that has been compromised because of target site modification. Resistant strains containing a single mutation in rRNA developed within seven months of its use in the clinical settings.

#### BRIEF SUMMARY OF THE INVENTION

The "instant evolution" system was initially developed in *E. coli*, primarily because of the ease with which this organism can be genetically manipulated. [WO 2004/003511.] Because many of the functionally important regions of rRNA are conserved among bacteria, drug leads developed against conserved targets in the *E. coli* system may produce broad-spectrum anti-infectives. However, in order to develop a system to produce narrow-spectrum anti-infectives, provided herein are methods and compositions for screening *Pseudomonas aeruginosa* 16S rRNA in *E. coli* cells. In certain embodiments, a plasmid comprising the 16S rRNA gene from *Pseudomonas aeruginosa*, mutated to replace the natural helix 9 region with the corresponding region of the *E. coli* rRNA, is provided to form functional ribosomes in *E. coli* host cells. In other embodiments, a plasmid, comprising the unmutated 16S rRNA from *Pseudomonas aeruginosa*, along with a plasmid containing

the *Pseudomonas aeruginosa* S20 protein, is provided which can yield functional ribosomes in *E. coli* cells.

#### BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 depicts the plasmid construct Pa 16S pRNA228 in the bottom panel, wherein the *E. coli* 16S is replaced with the *P. aeruginosa* between the *E. coli* BclI and EstEII sites. The locations of specific sites in Pa 16S pRNA228 are shown in the upper panel as follows: the 16S rRNA *E. coli* rrnB operon corresponds to nucleic acids 1-14 & 1499-1536; the 16S rRNA *P. aeruginosa* rrnC operon corresponds to nucleic acids 15-1498; the 16S MBS (message binding sequence) GGGAU corresponds to nucleic acids 1530-1534; the 16S-23S spacer region corresponds to nucleic acids 1537-1976; the 23S rRNA of *E. coli* rrnB operon corresponds to nucleic acids 1977-4880; the 23S-5S spacer region corresponds to nucleic acids 4881-4972; the 5S rRNA of *E. coli* rrnB operon corresponds to nucleic acids 4973-5092; the terminator T1 of *E. coli* rrnB operon corresponds to nucleic acids 5096-5139; the terminator T2 of *E. coli* rrnB operon corresponds to nucleic acids 5271-5299; the bla ( $\beta$ -lactamase; ampicillin resistance) corresponds to nucleic acids 6462-7319; the replication origin corresponds to nucleic acids 7482-8096; the rop (Rop protein) corresponds to nucleic acids 8509-8700; the GFP corresponds to nucleic acids 9113-10088; the GFP RBS (ribosome binding sequence) AUCCC corresponds to nucleic acids 10096-10100; the trp<sup>c</sup> promoter corresponds to nucleic acids 10117-10157; the trp<sup>c</sup> promoter corresponds to nucleic acids 10632-10672; the CAT RBS AUCCC corresponds to nucleic acids 10689-10693; the cam (chloramphenicol acetyltransferase: CAT) corresponds to nucleic acids 10701-11630; the lacY<sup>1</sup> promoter corresponds to nucleic acids 11670-11747; the /ac/q (lac repressor) corresponds to nucleic acids 11748-12830; and the lacUV5 promoter corresponds to nucleic acids 12873-12914.

FIG. 2A-B depict the plasmid construct Pa 16S Ec H9 pRNA228. In FIG. 2B, the plasmid map is shown wherein the *E. coli* 16S is replaced with the *P. aeruginosa* between the *E. coli* BclI and EstEII sites and the *P. aeruginosa* helix 9 (H9) sequence is replaced with the *E. coli* H9 sequence. FIG. 2A shows that the plasmid is made up of *E. coli* rRNA sequence at positions 1-14, 176-188 and 1499-1536; and *P. aeruginosa* rRNA sequence at positions 15-175 and 189-1489. The *E. coli* H9 sequence is AUAACGUCGCAAGAC-CAAA (SEQ ID NO: 1) and the *P. aeruginosa* H9 sequence is

(SEQ ID NO: 2)

AUACGUCCUGAGGGAGAAA.

10 mutations were made and they are denoted with an underline in the *P. aeruginosa* H9 sequence shown above.

FIG. 3 depicts a pKan5-T1T2 vector in the bottom panel wherein the *E. coli* terminators from pRNA228 were moved (via PCR) into the multicloning site of pKANS using restriction enzymes XbaI and EcoRI. The locations of specific sites in pKan5-T1T2 are shown in the upper panel as follows: the replication origin corresponds to nucleic acids 723-1268; aph (3')-Ia (kanamycin resistance) corresponds to nucleic acids 1862-2677; araC corresponds to nucleic acids 2950-3828; Pc (the araC promoter) corresponds to nucleic acids 3979-4005; the pBAD promoter corresponds to nucleic acids 4104-4131; the terminator T1 of *E. coli* rrnB

operon corresponds to nucleic acids 4163-4206; and the terminator T2 of *E. coli* *rrnB* operon corresponds to nucleic acids 4338-4366.

FIG. 4 depicts a pKan5 vector in the bottom panel wherein the the *E. coli* terminators from pRNA228 were moved (via PCR) into the multicloning site of pKan5 using restriction enzymes XbaI and EcoRI; and the *P. aeruginosa* S20 protein was also cloned into the vector using enzymes NotI and XbaI (denoted herein as "pKanPa-S20"). The locations of specific sites in pKan5Pa-S20 are shown in the upper panel as follows: the replication origin corresponds to nucleic acids 723-1268; aph (3)-Ia (kanamycin resistance) corresponds to nucleic acids 1862-2677; araC corresponds to nucleic acids 2950-3828; Pc (the araC promoter) corresponds to nucleic acids 3979-4005; the pBAD promoter corresponds to nucleic acids 4104-4131; the Shine-Dalgarno sequence (GAGGA) corresponds to nucleic acids 4168-4172; *P. aeruginosa* S20 corresponds to 4181-4453; the terminator T1 of *E. coli* *rrnB* operon corresponds to nucleic acids 4463-4506; and the terminator T2 of *E. coli* *rrnB* operon corresponds to nucleic acids 4638-4666.

FIG. 5 depicts the results of green fluorescent protein (GFP) assays for *P. aeruginosa* constructs.

FIGS. 6a-f depict the sequence of the plasmid depicted in FIG. 1 (SEQ ID NO: 3).

FIGS. 7a-f depict the sequence of the plasmid depicted in FIG. 2 (SEQ ID NO: 4).

FIGS. 8a-b depict the sequence of the plasmid depicted in FIG. 3 (SEQ ID NO: 5).

FIGS. 9a-b depict the sequence of the plasmid depicted in FIG. 4 (SEQ ID NO: 6).

FIG. 10 depicts novel mutant anti-Shine-Dalgarno (ASD) sequences (SEQ ID NOs: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30 and 32) and novel mutant Shine-Dalgarno (SD) sequences (SEQ ID NOs: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, and 31) of the present invention. FIG. 10 also shows a sequence analysis of chloramphenicol resistant isolates. The mutated nucleotides are underlined and potential duplex formations are boxed. CAT activity was measured twice for each culture and the unit is CPM/0.1  $\mu$ L of culture/OD600. Induction was measured by dividing CAT activity in induced cells with CAT activity in uninduced cells. A -1 indicates no induction, while a +1 indicates induction with 1 mM IPTG.

FIG. 11 depicts novel mutant ASD sequences (SEQ ID NOs: 34, 36, 38, 40, 42, 44, 46 and 48) and novel mutant SD sequences (SEQ ID NOs: 33, 35, 37, 39, 41, 43, 45 and 47), of the present invention. FIG. 11 also shows a sequence analysis of CAT mRNA mutants (SEQ ID NOs: 34, 36, 38, 40, 42, 44, 46 and 48). Potential duplex formations are boxed and the mutated nucleotides are underlined. The start codon (AUG) is in bold. -1 in the table indicates no induction, whereas +1 indicates induction with 1 mM IPTG.

FIGS. 12a-b depict novel nucleic acid mutant ASD sequences and novel complementary mutant SD sequences. FIG. 12A-1 depicts novel mutant ASD sequences SEQ ID NOs: 50, 52, 54, 56 and 58 and novel mutant SD sequences SEQ ID NOs: 49, 51, 53, 55, and 57; FIG. 12A-2 depicts novel mutant ASD sequences SEQ ID NOs: 60, 62, 64, 66, 68, 70, 72, 74, 76, 78 and 80 and novel mutant SD sequences SEQ ID NOs: 59, 61, 63, 65, 67, 69, 71, 73, 75, 77 and 79, FIG. 12A-3 depicts novel mutant ASD sequences 82, 84, 86, 88, 90, 92, 94, 96, 98 and 100 and novel mutant SD sequences SEQ ID NOs: 81, 83, 85, 87, 89, 91, 93, 95, 97 and 99; FIG. 12B-1 depicts novel mutant ASD sequences SEQ ID NOs: 102, 104, and 106 and novel mutant SD sequences SEQ ID NOs: 101, 103 and 105; FIG. 12B-2 depicts novel ASD sequences SEQ ID NOs: 108, 110, 112,

114, 116, 118, 120, 122, 124, 126, and 128 and novel SD sequences SEQ ID NOs: 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, and 127, and FIG. 12B-3 depicts novel ASD sequences SEQ ID NOs: 130, 132, 134, 136, 138, 140, 142, 144, 146 and 148 and novel SD sequences SEQ ID NOs: 129, 131, 133, 135, 137, 139, 141, 143, 145 and 147.

FIG. 13a-i depict *P. aeruginosa* 16S rRNA (SEQ ID NO: 149), with the shown nucleotide differences between *E. coli* (SEQ ID NO: 150) and *P. aeruginosa* 16S rRNA (SEQ ID NO: 149). The *P. aeruginosa* 16S rRNA is shown in full, with the *E. coli* sequence differences shown in circles.

FIG. 14 depicts a tabulation of 16s rRNA hybrid *P. aeruginosa* and *E. coli* sequences and the resulting measured GFP percentage. (GFP 100% was set for *E. coli* sequence with no *P. aeruginosa* sequence.) The last entry in the table demonstrates the significance of nucleotides 176-188; this is shown pictorially in FIG. 15.

FIG. 15a-h depict a mutant *P. aeruginosa* 16S rRNA wherein the helix 9 (H9) nucleotides, numbered 176-188 (boxed in the lower left) have been replaced with the corresponding *E. coli* H9 nucleotides (SEQ ID NO: 151).

FIG. 16 depicts the ribosomal protein S20 binding site on *P. aeruginosa* 16S rRNA (SEQ ID NO: 152).

FIG. 17 depicts an S20 protein alignment of *E. coli* (SEQ ID NO: 154) compared with *P. aeruginosa* (SEQ ID NO: 153) and *Thermus* species (SEQ ID NO: 155).

FIG. 18 depicts a schematic of the addition of a plasmid containing the S20 gene into transformed in cells containing the 16S rRNA of *E. coli* and *P. aeruginosa*, and the measurement of resulting activity by GFP.

In addition, FIGS. 1-26 of WO 2004/003511 (some of which may also be found in Lee, K., et al. Genetic Approaches to Studying Protein Synthesis: Effects of Mutations at Pseudouridine 516 and A535 in *Escherichia coli* 16S rRNA. Symposium: Translational Control: A Mechanistic Perspective at the Experimental Biology 2001 Meeting (2001); and Lee, K., et al., J. Mol. Biol. 269: 732-743 (1997)) are expressly incorporated herein by reference.

#### DETAILED DESCRIPTION OF THE INVENTION

Overview. The "instant evolution" system (as described in WO 2004/003511; herein incorporated by reference in its entirety) was initially developed in *E. coli*, primarily because of the ease with which this organism can be genetically manipulated. Because many of the functionally important regions of rRNA are conserved among bacteria, drug leads developed against conserved targets in the *E. coli* system may produce broad-spectrum anti-infectives.

Recently, however, emphasis has been placed on the development of narrow spectrum antimicrobials because of the reduced likeliness of adverse effects and because resistance to narrow-spectrum antibiotics is less likely to occur. Several rRNA genes from human pathogens were therefore tested to see whether their genes could be expressed in *E. coli*.

Compositions and Methods of the Invention. In one embodiment, compositions and methods are provided to identify functional mutant *P. aeruginosa* ribosomes suitable as drug targets. The compositions and methods of the invention allow for the expression of *P. aeruginosa* ribosomes in *E. coli*. The compositions and methods further allow for the isolation and analysis of mutations that would normally be lethal and allow direct selection of rRNA mutants with predetermined levels of ribosome function. The compositions and methods of the present invention may

additionally be used to identify antibiotics to treat generally and/or selectively, human pathogens such as *P. aeruginosa*.

According to one embodiment of the invention, a functional genomics database for 16S rRNA genes of a variety of species is generated. In particular, the rRNA gene is randomly mutated using a generalized mutational strategy. A host *E. coli* cell is then transformed with a mutagenized plasmid of the invention comprising: an *P. aeruginosa* rRNA gene having a mutant ASD sequence, the mutated *P. aeruginosa* rRNA gene, and a genetically engineered gene which encodes a selectable marker having a mutant SD sequence. In certain embodiments, said *P. aeruginosa* DNA is further modified so as to replace helix 9 with the corresponding *E. coli* helix (FIG. 15); in other embodiments the host cell is further transformed with a plasmid encoding the *P. aeruginosa* S20 gene. The selectable marker gene, such as CAT, may be used to select mutants that are functional, e.g., by plating the transformed cells onto growth medium containing chloramphenicol. The mutant rRNA genes contained in each plasmid DNA of the individual clones from each colony are selected and characterized. The function of each of the mutant rRNA genes is assessed by measuring the amount of an additional selectable marker gene, such as GFP, produced by each clone upon induction of the rRNA operon. A functional genomics database may thus be assembled, which contains the sequence and functional data of the functional mutant rRNA genes. In particular, functionally important regions of the rRNA gene that will serve as drug targets are identified by comparing the sequences of the functional genomics database and correlating the sequence with the amount of GFP protein produced.

In another embodiment, the nucleotides in the functionally important target regions identified in the above methods may be simultaneously randomly mutated, e.g., by using standard methods of molecular mutagenesis, and cloned into a plasmid of the invention to form a plasmid pool containing random mutations at each of the nucleotide positions in the target region. The resulting pool of plasmids containing random mutations is then used to transform cells, e.g., *E. coli* cells, and form a library of clones, each of which contains a unique combination of mutations in the target region. The library of mutant clones are grown in the presence of IPTG to induce production of the mutant rRNA genes and a selectable marker is used, such as CAT, to select clones of rRNA mutants containing nucleotide combinations of the target region that produce functional ribosomes. The rRNA genes producing functional ribosomes are sequenced and may be incorporated into a database.

In yet another embodiment, a series of oligonucleotides may be synthesized that contain the functionally-important nucleotides and nucleotide motifs within the target region and may be used to sequentially screen compounds and compound libraries to identify compounds that recognize (bind to) the functionally important sequences and motifs. The compounds that bind to all of the oligonucleotides are then counter-screened against oligonucleotides and/or other RNA containing molecules to identify drug candidates. Drug candidates selected by the methods of the present invention are thus capable of recognizing all of the functional variants of the target sequence, i.e., the target cannot be mutated in a way that the drug cannot bind, without causing loss of function to the ribosome.

In still another embodiment, after the first stage mutagenesis of the entire rRNA is performed using techniques known in the art, e.g., error-prone PCR mutagenesis, the mutants are analyzed to identify regions within the rRNA that are important for function. These regions are then sorted

based on their phylogenetic conservation, as described herein, and are then used for further mutagenesis.

Ribosomal RNA sequences from each species are different and the more closely related two species are, the more their rRNAs are alike. For instance, humans and monkeys have very similar rRNA sequences, but humans and bacteria have very different rRNA sequences. These differences may be utilized for the development of very specific drugs with a narrow spectrum of action and also for the development of broad-spectrum drugs that inhibit large groups of organisms that are only distantly related, such as all bacteria.

In another embodiment, the functionally important regions identified above are divided into groups based upon whether or not they occur in closely related groups of organisms. For instance, some regions of rRNA are found in all bacteria but not in other organisms. Other areas of rRNA are found only in closely related groups of bacteria, such as all of the members of a particular species, e.g., members of the genus *Mycobacterium* or *Streptococcus*.

In a further embodiment, the regions found in very large groups of organisms, e.g., all bacteria or all fungi, are used to develop broad-spectrum antibiotics that may be used to treat infections from a large number of organisms within that group. The methods of the present invention may be performed on these regions and functional mutant ribosomes identified. These functional mutant ribosomes may be screened, for example, with compound libraries.

In yet another embodiment, regions that are located only in relatively small groups of organisms, such as all members of the genus *Streptococcus* or all members of the genus *Mycobacterium*, may be used to design narrow spectrum antibiotics that will only inhibit the growth of organisms that fall within these smaller groups. The methods of the present invention may be performed on these regions and functional mutant ribosomes identified. These functional mutant ribosomes will be screened, e.g., compound libraries.

The invention provides novel plasmid constructs, e.g. those depicted in FIGS. 1 to 4. In addition, the novel plasmid constructs of the present invention employ novel mutant anti-Shrine-Delgano (ASD) and mutant Shine-Delgano (SD) sequences set forth in FIGS. 10, 11 and 12. The mutant ASD and mutant SD sequences may be used as mutually compatible pairs (see FIGS. 10, 11 and 12). It will be appreciated that the mutually compatible pairs of mutant ASD and SD sequences interact as pairs in the form of RNA, to permit translation of only the mRNAs containing the altered SD sequence.

Definitions. For convenience, certain terms employed in the specification, examples, and appended claims are collected here.

The articles "a" and "an" are used herein to refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

An "inducible" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living cell substantially only when an inducer which corresponds to the promoter is present in the cell.

As used herein, the term "mutation" includes an alteration in the nucleotide sequence of a given gene or regulatory sequence from the naturally occurring or normal nucleotide sequence. A mutation may be a single nucleotide alteration (e.g., deletion, insertion, substitution, including a point mutation), or a deletion, insertion, or substitution of a number of nucleotides.

By the term "selectable marker" is meant a gene whose expression allows one to identify functional mutant ribosomes. Useful selectable markers are well known in the art and include, for example, antibiotic and herbicide resistance genes, gas concentrations, nutrients, and waste products. Specific examples of such genes are disclosed in Weising et al. Weising, K, et al., (1988) Ann Rev of Genetics 22:421-478; the contents of which are incorporated by reference. Examples of selectable marker genes are genes encoding a protein such as G418 and hygromycin which confer resistance to certain drugs, for example,  $\beta$ -galactosidase, chloramphenicol acetyltransferase, or firefly luciferase. Transcription of the selectable marker gene is monitored by changes in the concentration of the selectable marker protein such as  $\beta$ -galactosidase, chloramphenicol acetyltransferase, or firefly luciferase. If the selectable marker gene encodes a protein conferring antibiotic resistance such as neomycin resistance transformant cells can be selected with G418. Cells that have incorporated the selectable marker gene will survive, while the other cells die. This makes it possible to visualize and assay for expression of recombinant expression vectors of the invention and in particular to determine the effect of a mutation on expression and phenotype. It will be appreciated that selectable markers can be introduced on a separate vector from the nucleic acid of interest.

The term 'Gram-positive bacteria' is an art recognized term for bacteria characterized by having as part of their cell wall structure peptidoglycan as well as polysaccharides and/or teichoic acids and are characterized by their blue-violet color reaction in the Gram-staining procedure. Representative Gram-positive bacteria include: *Actinomyces* spp., *Bacillus anthracis*, *Bifidobacterium* spp., *Clostridium botulinum*, *Clostridium perfringens*, *Clostridium* spp., *Clostridium tetani*, *Corynebacterium diphtheriae*, *Corynebacterium jeikeium*, *Enterococcus faecalis*, *Enterococcus faecium*, *Erysipelothrix rhusiopathiae*, *Eubacterium* spp., *Gardnerella vaginalis*, *Gemella morbillorum*, *Leuconostoc* spp., *Mycobacterium abscessus*, *Mycobacterium avium* complex, *Mycobacterium chelonae*, *Mycobacterium fortuitum*, *Mycobacterium haemophilium*, *Mycobacterium kansasii*, *Mycobacterium leprae*, *Mycobacterium marinum*, *Mycobacterium scrofulaceum*, *Mycobacterium smegmatis*, *Mycobacterium terrae*, *Mycobacterium tuberculosis*, *Mycobacterium ulcerans*, *Nocardia* spp., *Peptococcus niger*, *Peptostreptococcus* spp., *Propionibacterium* spp., *Staphylococcus aureus*, *Staphylococcus auricularis*, *Staphylococcus capitis*, *Staphylococcus cohnii*, *Staphylococcus epidermidis*, *Staphylococcus haemolyticus*, *Staphylococcus hominis*, *Staphylococcus lugdanensis*, *Staphylococcus saccharolyticus*, *Staphylococcus saprophyticus*, *Staphylococcus schleiferi*, *Staphylococcus similans*, *Staphylococcus warneri*, *Staphylococcus xylosus*, *Streptococcus agalactiae* (group B streptococcus), *Streptococcus anginosus*, *Streptococcus bovis*, *Streptococcus canis*, *Streptococcus equi*, *Streptococcus milleri*, *Streptococcus mitior*, *Streptococcus mutans*, *Streptococcus pneumoniae*, *Streptococcus pyogenes* (group A streptococcus), *Streptococcus salivarius*, *Streptococcus sanguis*.

The term "Gram-negative bacteria" is an art recognized term for bacteria characterized by the presence of a double membrane surrounding each bacterial cell. Representative Gram-negative bacteria include *Acinetobacter calcoaceticus*, *Actinobacillus actinomycetemcomitans*, *Aeromonas hydrophila*, *Alcaligenes xylosoxidans*, *Bacteroides*, *Bacteroides fragilis*, *Bartonella bacilliformis*, *Bordetella* spp., *Borrelia burgdorferi*, *Branhamella catarrhalis*, *Brucella* spp., *Campylobacter* spp., *Chlamydia pneumoniae*, *Chla-*

*mydia psittaci*, *Chlamydia trachomatis*, *Chromobacterium violaceum*, *Citrobacter* spp., *Eikenella corrodens*, *Enterobacter aerogenes*, *Escherichia coli*, *Flavobacterium meningosepticum*, *Fusobacterium* spp., *Haemophilus influenzae*, *Haemophilus* spp., *Helicobacter pylori*, *Klebsiella* spp., *Legionella* spp., *Leptospira* spp., *Moraxella catarrhalis*, *Morganella morganii*, *Mycoplasma pneumoniae*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Pasteurella multocida*, *Plesiomonas shigelloides*, *Prevotella* spp., *Proteus* spp., *Providencia rettgeri*, *Pseudomonas aeruginosa*, *Pseudomonas* spp., *Rickettsia prowazekii*, *Rickettsia rickettsii*, *Rochalimaea* spp., *Salmonella* spp., *Salmonella typhi*, *Serratia marcescens*, *Shigella* spp., *Treponema carateum*, *Treponema pallidum*, *Treponema pallidum endemicum*, *Treponema pertenuis*, *Veillonella* spp., *Vibrio cholerae*, *Vibrio vulnificus*, *Yersinia enterocolitica*, *Yersinia pestis*.

Isolated Nucleic Acid Molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium, when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence set forth in FIGS. 10, 11, and 12, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence set forth in FIGS. 10, 11, and 12 as a hybridization probe, the nucleic acid molecules of the present invention can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of the sequence set forth in FIGS. 10, 11, and 12, can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence set forth in FIGS. 10, 11, and 12.

A nucleic acid of the invention can be amplified using cDNA, mRNA or, alternatively, genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to the nucleotide sequences of the present invention can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.



In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence set forth in FIGS. 10, 11, and 12, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in FIGS. 10, 11, and 12 is one which is sufficiently complementary to the nucleotide sequence shown in FIGS. 10, 11, and 12, such that it can hybridize to the nucleotide sequence shown in FIGS. 10, 11, and 12, respectively, thereby forming a stable duplex.

“Homology” or alternatively “identity” refers to sequence similarity between two nucleic acid molecules. Homology may be determined by comparing a position in each sequence, which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences.

The term “percent identical” refers to sequence identity between two nucleotide sequences. Identity may be determined by comparing a position in each sequence, which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site is occupied by the same or a similar base (e.g., similar in steric and/or electronic nature), then the molecules may be referred to as homologous (similar) at that position. Expression as a percentage of homology, similarity, or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. Various alignment algorithms and/or programs may be used, including FASTA, BLAST, or ENTREZ. FASTA and BLAST are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and may be used with, e.g., default settings. ENTREZ is available through the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Md. In one embodiment, the percent identity of two sequences may be determined by the GCG program with a gap weight of 1, e.g., each nucleic acid gap is weighted as if it were a single nucleotide mismatch between the two sequences. Other techniques for alignment are described in *Methods in Enzymology*, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, Calif., USA. Preferably, an alignment program that permits gaps in the sequence is utilized to align the sequences. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See *Meth. Mol. Biol.* 70: 173-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method may be utilized to align sequences. An alternative search strategy uses MPSRCH software, which runs on a MASPAR computer. MPSRCH uses a Smith-Waterman algorithm to score sequences on a massively parallel computer. This approach improves the ability to pick up distantly related matches, and is especially tolerant of small gaps and nucleotide sequence errors. Nucleic acid-encoded amino acid sequences may be used to search both protein and DNA databases. Databases with individual sequences are described in *Methods in Enzymology*, ed. Doolittle, supra. Databases include Genbank, EMBL, and DNA Database of Japan (DDBJ).

The terms “polynucleotide”, and “nucleic acid” are used interchangeably to refer to a polymeric form of nucleotides

of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. The term “recombinant” polynucleotide means a polynucleotide of genomic, cDNA, semi-synthetic, or synthetic origin, which either does not occur in nature or is linked to another polynucleotide in a non-natural arrangement. An “oligonucleotide” refers to a single stranded polynucleotide having less than about 100 nucleotides, less than about, e.g., 75, 50, 25, or 10 nucleotides.

The term “specifically hybridizes” refers to detectable and specific nucleic acid binding. Polynucleotides, oligonucleotides and nucleic acids of the invention selectively hybridize to nucleic acid strands under hybridization and wash conditions that minimize appreciable amounts of detectable binding to nonspecific nucleic acids. Stringent conditions may be used to achieve selective hybridization conditions as known in the art and discussed herein. Generally, the nucleic acid sequence homology between the polynucleotides, oligonucleotides, and nucleic acids of the invention and a nucleic acid sequence of interest will be at least 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99%, or more. In certain instances, hybridization and washing conditions are performed under stringent conditions according to conventional hybridization procedures and as described further herein.

The terms “stringent conditions” or “stringent hybridization conditions” refer to conditions, which promote specific hybridization between two complementary polynucleotide strands so as to form a duplex. Stringent conditions may be selected to be about 5° C. lower than the thermal melting point (T<sub>m</sub>) for a given polynucleotide duplex at a defined ionic strength and pH. The length of the complementary polynucleotide strands and their GC content will determine the T<sub>m</sub> of the duplex, and thus the hybridization conditions necessary for obtaining a desired specificity of hybridization. The T<sub>m</sub> is the temperature (under defined ionic strength and pH) at which 50% of a polynucleotide sequence hybridizes to a perfectly matched complementary strand. In certain cases it may be desirable to increase the stringency of the hybridization conditions to be about equal to the T<sub>m</sub> for a particular duplex.

A variety of techniques for estimating the T<sub>m</sub> are available. Typically, G-C base pairs in a duplex are estimated to contribute about 3° C. to the T<sub>m</sub>, while A-T base pairs are estimated to contribute about 2° C., up to a theoretical maximum of about 80-100° C. However, more sophisticated models of T<sub>m</sub> are available in which G-C stacking interactions, solvent effects, the desired assay temperature and the like are taken into account. For example, probes can be designed to have a dissociation temperature (T<sub>d</sub>) of approximately 60° C., using the formula:  $T_d = (((((3 \times \#GC) + (2 \times \#AT)) \times 37) - 562) / \#bp) - 5$ ; where #GC, #AT, and #bp are the number of guanine-cytosine base pairs, the number of

adenine-thymine base pairs, and the number of total base pairs, respectively, involved in the formation of the duplex.

Hybridization may be carried out in SxSSC, 4xSSC, 3xSSC, 2xSSC, 1xSSC or 0.2xSSC for at least about 1 hour, 2 hours, 5 hours, 12 hours, or 24 hours. The temperature of the hybridization may be increased to adjust the stringency of the reaction, for example, from about 25° C. (room temperature), to about 45° C., 50° C., 55° C., 60° C., or 65° C. The hybridization reaction may also include another agent affecting the stringency, for example, hybridization conducted in the presence of 50% formamide increases the stringency of hybridization at a defined temperature.

The hybridization reaction may be followed by a single wash step, or two or more wash steps, which may be at the same or a different salinity and temperature. For example, the temperature of the wash may be increased to adjust the stringency from about 25° C. (room temperature), to about 45° C., 50° C., 55° C., 60° C., 65° C., or higher. The wash step may be conducted in the presence of a detergent, e.g., 0.1 or 0.2% SDS. For example, hybridization may be followed by two wash steps at 65° C. each for about 20 minutes in 2xSSC, 0.1% SDS, and optionally two additional wash steps at 65° C. each for about 20 minutes in 0.2xSSC, 0.1% SDS.

Exemplary stringent hybridization conditions include overnight hybridization at 65° C. in a solution comprising, or consisting of, 50% formamide, 10xDenhardt (0.2% Ficoll, 0.2% Polyvinylpyrrolidone, 0.2% bovine serum albumin) and 200 µg/ml of denatured carrier DNA, e.g., sheared salmon sperm DNA, followed by two wash steps at 65° C. each for about 20 minutes in 2xSSC, 0.1% SDS, and two wash steps at 65° C. each for about 20 minutes in 0.2xSSC, 0.1% SDS.

Hybridization may consist of hybridizing two nucleic acids in solution, or a nucleic acid in solution to a nucleic acid attached to a solid support, e.g., a filter. When one nucleic acid is on a solid support, a prehybridization step may be conducted prior to hybridization. Prehybridization may be carried out for at least about 1 hour, 3 hours or 10 hours in the same solution and at the same temperature as the hybridization solution (without the complementary polynucleotide strand).

Appropriate stringency conditions are known to those skilled in the art or may be determined experimentally by the skilled artisan. See, for example, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-12.3.6; Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; S. Agrawal (ed.) Methods in Molecular Biology, volume 20; Tijssen (1993) Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes, e.g., part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York; and Tibanyenda, N. et al., Eur. J. Biochem. 139:19 (1984) and Ebel, S. et al., Biochem. 31:12083 (1992).

The term "substantially homologous" when used in connection with a nucleic acid or amino acid sequences, refers to sequences which are substantially identical to or similar in sequence with each other, giving rise to a homology of conformation and thus to retention, to a useful degree, of one or more biological (including immunological) activities. The term is not intended to imply a common evolution of the sequences.

Recombinant Expression Vectors and Host Cells. Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid molecule of the present invention (or a portion thereof). As used herein, the

term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel (1990) Methods Enzymol. 185:3-7. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety

subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D. B. and Johnson, K. S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRITS (Pharmacia, Piscataway, N.J.) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al. (1988) *Gene* 69:301-315) and pET lld (Studier et al. (1990) *Methods Enzymol.* 185:60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET lld vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gnl). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gnl gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S. (1990) *Methods Enzymol.* 185:119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the expression vector may be a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari et al. (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz (1982) *Cell* 30:933-943), pJRY88 (Schultz et al. (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J. et al., *Molecular Cloning: A Laboratory Manual*. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-

specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example by the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379).

Another aspect of the invention pertains to host cells into which a the nucleic acid molecule of the invention is introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook et al. (*Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

Making *P. aeruginosa* 16S rRNA active in *E. coli*. As already mentioned herein, a plasmid comprising the 16S rRNA gene from *P. aeruginosa*, mutated to replace the natural helix 9 region with the corresponding region of the *E. coli* rRNA, is provided to form functional mutant ribosomes. In addition, in certain embodiments, a plasmid comprising the *Pseudomonas aeruginosa* S20 protein is provided, thereby yielding functional ribosomes in *E. coli* cells.

There are many different ribosomal proteins. The 30S ribosomal proteins are designated S1-S21. It is known that S4, S7, S8, S15, S17 and S20 bind independently to 16S rRNA. Binding of these primary binding proteins folds the rRNA and allows binding of the secondary binding proteins (S5, S6, S9, S12, S13, S16, S18, and S19). These proteins also fold the ribosome to allow the addition of S2, S3, S10, S11, S14 and S21 (which are known as tertiary binding proteins).

The ribosomal protein S20 binding site on *P. aeruginosa* are shown in FIG. 16. 16S rRNA Protein alignments for *E. coli* and *P. aeruginosa* are shown in FIG. 17. As shown therein, S20's pH2 and pH3 interact with H9 of the 16S rRNA. It is hypothesized that charge differences between *E. coli* and *P. aeruginosa* S20 composition may play a role in binding of the protein to the rRNA.

The expression of *P. aeruginosa* S20 can be made in the presence of *E. coli* and *P. aeruginosa* 16S RNA. One begins by obtaining chromosomal DNA from *P. aeruginosa*. Then the S20 gene is PCR'd. Next, this gene is cloned into an expression vector. Finally, the plasmid containing the S20 gene is transformed in cells (e.g. DH5 cells) containing the 16S rRNA of *E. coli* and *P. aeruginosa*, and the activity is measured. A schematic of this process is shown in FIG. 18. GFP analysis of *P. aeruginosa* (46% activity of ribosomes without *P. aeruginosa* S20 and 96.4% activity of ribosomes with *P. aeruginosa* S20) confirms that *P. aeruginosa* 16S rRNA is complemented by *P. aeruginosa* S20. When the vector only was assayed (i.e. no S20) no increase in ribosomal activity was observed. *P. aeruginosa* 16S rRNAs may therefore be made active in *E. coli* by replacing helix 9 of *P. aeruginosa* 16S rRNA with the corresponding sequence from *E. coli* or by adding *P. aeruginosa* S20 to cells that are making the *P. aeruginosa* 16S rRNA. Both constructs can be used to develop antibiotics that target *P. aeruginosa* 16S rRNA.

Making other pathogen's 16S rRNA active in *E. coli*. Given *P. aeruginosa* 16S rRNA can be fully expressed in *E. coli*, it is likely that the rRNAs from other pathogens (e.g. Gram-positive or Gram-negative bacteria) can also be expressed in *E. coli*. This will facilitate the identification of new antibiotic targets, and therefore new classes of antibiotics which are less susceptible to the development of resistance. The inventive system described herein can be used to identify antibiotic leads that disrupt the interaction between critical nucleotides of pathogenic 16S rRNA and the critical amino acids on ribosomal proteins (such as critical nucleotides on *P. aeruginosa* 16S rRNA and *P. aeruginosa* S20). An antibiotic developed in this manner with specifically target the pathogen used, and its close relatives, without effecting other bacteria.

Uses and Methods of the Invention. The nucleic acid molecules described herein may be used in a plasmid construct, e.g. pRNA228, to carry out one or more of the following methods: (1) creation of a functional genomics database of the rRNA genes generated by the methods of the present invention; (2) mining of the database to identify functionally important regions of the rRNA; (3) identification of functionally important sequences and structural motifs within each target region; (4) screening compounds and compound libraries against a series of functional variants of the target sequence to identify compounds that bind to all functional variants of the target sequence; and (5) counterscreening the compounds against nontarget RNAs, such as human ribosomes or ribosomal RNA sequences.

One aspect of the invention relates to a plasmid comprising a first nucleic acid sequence and a second nucleic acid sequence; wherein said first nucleic acid sequence encodes a *Pseudomonas aeruginosa* 16S rRNA comprising a mutant Anti-Shine-Dalgarno sequence and at least one additional mutation outside the Anti-Shine-Dalgarno region; and said second sequence encodes a selectable marker having a mutant Shine-Dalgarno sequence; wherein said mutant Anti-Shine-Dalgarno specifically hybridizes to said mutant Shine-Dalgarno sequence.

Another aspect of the invention relates to the plasmid of FIG. 1.

Another aspect of the invention relates to a plasmid comprising a first nucleic acid sequence and a second nucleic acid sequence; wherein said first nucleic acid sequence encodes a *Pseudomonas aeruginosa* 16S rRNA comprising a mutant Anti-Shine-Dalgarno sequence, a mutant helix 9 sequence, and at least one additional mutation outside the Anti-Shine-Dalgarno region; and said second sequence encodes a selectable marker having a mutant Shine-Dalgarno sequence; wherein said mutant Anti-Shine-Dalgarno specifically hybridizes to said mutant Shine-Dalgarno sequence.

Another aspect of the invention relates to the plasmid of FIG. 2.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the selectable marker is chosen from the group consisting of chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), and both CAT and GFP.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein said mutant helix 9 sequence is SEQ ID NO: 1.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the mutant Anti-Shine-Dalgarno sequence is selected from the group consisting of the sequences set forth in FIGS. 10, 11 and 12.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the mutant Shine-Dalgarno sequence is selected from the group consisting of the sequences set forth in FIGS. 10, 11 and 12.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the mutant Anti-Shine-Dalgarno sequence and the mutant SD sequence are a complementary pair selected from the group consisting of the sequences set forth in FIGS. 10, 11 and 12.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the complementary mutant Shine-Dalgarno and mutant Anti-Shine-Dalgarno pair permits translation by the rRNA of the selectable marker.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the selectable marker is CAT.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the selectable marker is GFP.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the DNA sequence encoding the rRNA gene is under the control of an inducible promoter.

Another aspect of the invention relates to a plasmid comprising a *Pseudomonas aeruginosa* S20 gene.

Another aspect of the invention relates to a plasmid of FIG. 3 or 4.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising aph (3')-Ia (kanamycin resistance) gene.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising the terminator T1 and terminator T2 of the *Escherichia coli* rrnA, rrnB, rrnC, rrnD, rrnE, rrnG or rrnH operon; or a T1 terminator from one and a T2 terminator from another.

In certain embodiments the present invention relates to the aforementioned plasmid, comprising an operon from a bacteria.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising the terminator T1 and terminator T2 of the *Escherichia coli* rrnB operon.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising the terminator T1 and terminator T2 of the *Escherichia coli* rrnC operon.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising the terminator T1 and terminator T2 of an *Escherichia coli* rrn operon.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising the pBAD promotor and the AraC activator.

In certain embodiments the present invention relates to the aforementioned plasmid, further comprising the Shine-Dalgarno sequence GAGGA.

In certain embodiments the present invention relates to the aforementioned plasmid, wherein the DNA sequence encoding the rRNA gene is under the control of an inducible promoter.

Another aspect of the invention relates to a cell comprising an aforementioned plasmid. In certain embodiments said cell comprises a plasmid which comprises a first nucleic acid sequence and a second nucleic acid sequence; wherein said first nucleic acid sequence encodes a *Pseudomonas aeruginosa* 16S rRNA comprising a mutant Anti-Shine-Dalgarno sequence and at least one additional mutation outside the Anti-Shine-Dalgarno region; and said second sequence encodes a selectable marker having a mutant Shine-Dalgarno sequence; wherein said mutant Anti-Shine-Dalgarno specifically hybridizes to said mutant Shine-Dalgarno sequence. In certain embodiments said cell comprises a plasmid which comprises a first nucleic acid sequence and a second nucleic acid sequence; wherein said first nucleic acid sequence encodes a *Pseudomonas aeruginosa* 16S rRNA comprising a mutant Anti-Shine-Dalgarno sequence and at least one additional mutation outside the Anti-Shine-Dalgarno region; and said second sequence encodes a selectable marker having a mutant Shine-Dalgarno sequence; wherein said mutant Anti-Shine-Dalgarno specifically hybridizes to said mutant Shine-Dalgarno sequence.

Another aspect of the invention relates to a cell comprising two of the aforementioned plasmids. In certain embodiments said cell comprises a plasmid which comprises a first nucleic acid sequence and a second nucleic acid sequence; wherein said first nucleic acid sequence encodes a *Pseudomonas aeruginosa* 16S rRNA comprising a mutant Anti-Shine-Dalgarno sequence and at least one additional mutation outside the Anti-Shine-Dalgarno region; and said second sequence encodes a selectable marker having a mutant Shine-Dalgarno sequence; wherein said mutant Anti-Shine-Dalgarno specifically hybridizes to said mutant Shine-Dalgarno sequence; and a plasmid which comprises a *Pseudomonas aeruginosa* S20 gene.

In certain embodiments the present invention relates to the aforementioned cell, wherein the mutations in the rRNA gene affect the quantity of selectable marker produced.

In certain embodiments the present invention relates to the aforementioned cell, wherein the cell is a bacterial cell.

In certain embodiments the present invention relates to the aforementioned cell, wherein the cell is an *E. coli* cell.

Another aspect of the present invention is a method for identifying functional mutant ribosomes comprising:

- (a) transforming a set of host cells with a set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene;
  - wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and
  - wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair;
  - thereby forming a set of transformed host cells;
- (b) transforming said set of host cells with a plasmid comprising a *Pseudomonas aeruginosa* S20 gene;
- (c) isolating from the set of transformed host cells those host cells which express the selectable marker gene product; and
- (d) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (c), thereby identifying functional mutant ribosomes.

Another aspect of the present invention is a method for identifying functional mutant ribosomes comprising:

- (a) transforming a set of host cells with a set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a green fluorescent protein gene;
  - wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation and a first mutant Anti-Shine-Dalgarno sequence; and said green fluorescent protein gene comprises a first mutant Shine-Dalgarno sequence; and
  - wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair;
  - thereby forming a set of transformed host cells;
- (b) transforming said set of host cells with a plasmid comprising a *Pseudomonas aeruginosa* S20 gene;
- (c) isolating from the set of transformed host cells those host cells which express the green fluorescent protein gene product; and
- (d) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (c), thereby identifying functional mutant ribosomes.

Another aspect of the present invention is a method for identifying functional mutant ribosomes that may be suitable as drug targets comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene;
  - wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and
  - wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair;
  - thereby forming a first set of transformed host cells;
- (b) transforming said first set of host cells with a plasmid comprising a *Pseudomonas aeruginosa* S20 gene;
- (c) isolating from the first set of transformed host cells those host cells which express the selectable marker gene product;

- (d) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (c) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced; 5
- (e) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (d) are mutated; and each rRNA gene further comprises a second mutant Anti-Shine-Dalgarno sequence; 10
- (f) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (e) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second selectable marker having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair; 20
- (g) transforming a second set of host cells with the plasmids from step (f) and plasmids comprising a *Pseudomonas aeruginosa* S20 gene, thereby forming a second set of transformed host cells;
- (h) isolating from the second set of transformed host cells from step (g) those host cells which express the selectable marker gene product; and 25
- (i) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (h), thereby identifying functional mutant ribosomes that may be suitable as drug targets. 30

Another aspect of the present invention is a method for identifying functional mutant ribosomes that may be suitable as drug targets comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a first green fluorescent protein gene; 35  
wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation and a first mutant Anti-Shine-Dalgarno sequence; and said first green fluorescent protein gene comprises a first mutant Shine-Dalgarno sequence; and 40  
wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; 45  
thereby forming a first set of transformed host cells;
- (b) transforming said first set of host cells with a plasmid comprising a *Pseudomonas aeruginosa* S20 gene;
- (c) isolating from the first set of transformed host cells those host cells which express the green fluorescent protein gene product; 50
- (d) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (c) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced; 55
- (e) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (d) are mutated; and each rRNA gene further comprises a second mutant Anti-Shine-Dalgarno sequence; 60
- (f) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (e) into a second set of plasmids; wherein said plasmids further comprise a 65

- second genetically engineered gene which encodes a second green fluorescent protein having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;
- (g) transforming a second set of host cells with the plasmids from step (f) and plasmids comprising a *Pseudomonas aeruginosa* S20 gene, thereby forming a second set of transformed host cells;
- (h) isolating from the second set of transformed host cells from step (g) those host cells which express the green fluorescent protein gene product; and
- (i) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (h), thereby identifying functional mutant ribosomes that may be suitable as drug targets.

Another aspect of the present invention is a method for identifying drug candidates comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene; 35  
wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and 40  
wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; 45  
thereby forming a first set of transformed host cells;
- (b) transforming said first set of host cells with a plasmid comprising a *Pseudomonas aeruginosa* S20 gene;
- (c) isolating from the first set of transformed host cells those host cells which express the selectable marker gene product;
- (d) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (c) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced; 50
- (e) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (d) are mutated; and each rRNA gene further comprises a second mutant Anti-Shine-Dalgarno sequence; 55
- (f) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (e) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second selectable marker having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair; 60
- (g) transforming a second set of host cells with the plasmids from step (f) and plasmids comprising a *Pseudomonas aeruginosa* S20 gene, thereby forming a second set of transformed host cells;
- (h) isolating from the second set of transformed host cells from step (g) those host cells which express the selectable marker gene product; 65
- (i) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (h), to identify the mutated regions of interest;

23

- (j) screening compounds against the mutated regions of interest from step (i) and wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (k) identifying the compounds from step (j) that bind to the mutated regions of interest from step (h) and the wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (l) screening the compounds from step (k) against human 16S rRNA; and
- (m) identifying the drug candidates from step (l) that do not bind to the human 16S rRNA, thereby identifying drug candidates.

Another aspect of the present invention is a method for identifying drug candidates comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a first green fluorescent protein gene; wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation and a first mutant Anti-Shine-Dalgarno sequence; and said first green fluorescent protein gene comprises a first mutant Shine-Dalgarno sequence; and wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; thereby forming a first set of transformed host cells;
- (b) transforming said first set of host cells with a plasmid comprising a *Pseudomonas aeruginosa* S20 gene;
- (c) isolating from the first set of transformed host cells those host cells which express the green fluorescent protein gene product;
- (d) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (c) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced;
- (e) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (d) are mutated; and each rRNA gene further comprises a second mutant Anti-Shine-Dalgarno sequence;
- (f) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (e) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second green fluorescent protein having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;
- (g) transforming a second set of host cells with the plasmids from step (f) and plasmids comprising a *Pseudomonas aeruginosa* S20 gene, thereby forming a second set of transformed host cells;
- (h) isolating from the second set of transformed host cells from step (g) those host cells which express the green fluorescent protein gene product;
- (i) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (h), to identify the mutated regions of interest;
- (j) screening compounds against the mutated regions of interest from step (i) and wildtype *Pseudomonas aeruginosa* 16S rRNA;

24

- (k) identifying the compounds from step (j) that bind to the mutated regions of interest from step (h) and the wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (l) screening the compounds from step (k) against human 16S rRNA; and
- (m) identifying the drug candidates from step (l) that do not bind to the human 16S rRNA, thereby identifying drug candidates.

Another aspect of the present invention is a method for identifying functional mutant ribosomes comprising:

- (a) transforming a set of host cells with a set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene; wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, a mutant helix 9 sequence, and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; thereby forming a set of transformed host cells;
- (b) isolating from the set of transformed host cells those host cells which express the selectable marker gene product; and
- (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b), thereby identifying functional mutant ribosomes.

Another aspect of the present invention is a method for identifying functional mutant ribosomes comprising:

- (a) transforming a set of host cells with a set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a green fluorescent protein gene; wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, a mutant helix 9 sequence, and a first mutant Anti-Shine-Dalgarno sequence; and said green fluorescent protein gene comprises a first mutant Shine-Dalgarno sequence; and wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; thereby forming a set of transformed host cells;
- (b) isolating from the set of transformed host cells those host cells which express the green fluorescent protein gene product; and
- (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b), thereby identifying functional mutant ribosomes.

Another aspect of the present invention is a method for identifying functional mutant ribosomes that may be suitable as drug targets comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene; wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, a mutant helix 9 sequence, and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and

- wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair;  
 thereby forming a first set of transformed host cells;  
 (b) isolating from the first set of transformed host cells those host cells which express the selectable marker gene product;  
 (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced;  
 (d) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (c) are mutated; and each mutant *Pseudomonas aeruginosa* 16S rRNA gene further comprises a mutant helix 9 sequence, and a second mutant Anti-Shine-Dalgarno sequence;  
 (e) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (d) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second selectable marker having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;  
 (f) transforming a second set of host cells with the plasmids from step (e), thereby forming a second set of transformed host cells;  
 (g) isolating from the second set of transformed host cells from step (f) those host cells which express the selectable marker gene product; and  
 (h) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (g), thereby identifying functional mutant ribosomes that may be suitable as drug targets.

Another aspect of the present invention is a method for identifying functional mutant ribosomes that may be suitable as drug targets comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a first green fluorescent protein gene;  
 wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, a mutant helix 9 sequence, and a first mutant Anti-Shine-Dalgarno sequence; and said first green fluorescent protein gene comprises a first mutant Shine-Dalgarno sequence; and  
 wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair;  
 thereby forming a first set of transformed host cells;  
 (b) isolating from the first set of transformed host cells those host cells which express the green fluorescent protein gene product;  
 (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced;  
 (d) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (c) are mutated; and each mutant

- Pseudomonas aeruginosa* 16S rRNA gene further comprises a mutant helix 9 sequence, and a second mutant Anti-Shine-Dalgarno sequence;  
 (e) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (d) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second green fluorescent protein having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;  
 (f) transforming a second set of host cells with the plasmids from step (e), thereby forming a second set of transformed host cells;  
 (g) isolating from the second set of transformed host cells from step (f) those host cells which express the green fluorescent protein gene product; and  
 (h) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (g), thereby identifying functional mutant ribosomes that may be suitable as drug targets.

Another aspect of the present invention is a method for identifying drug candidates comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene;  
 wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, a mutant helix 9 sequence, and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and  
 wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair;  
 thereby forming a first set of transformed host cells;  
 (b) isolating from the first set of transformed host cells those host cells which express the selectable marker gene product;  
 (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced;  
 (d) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (c) are mutated; and each mutant *Pseudomonas aeruginosa* 16S rRNA gene further comprises a mutant helix 9 sequence, and a second mutant Anti-Shine-Dalgarno sequence;  
 (e) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (d) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second selectable marker having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;  
 (f) transforming a second set of host cells with the plasmids from step (e), thereby forming a second set of transformed host cells;



27

- (g) isolating from the second set of transformed host cells from step (f) those host cells which express the selectable marker gene product;
- (h) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (g), to identify the mutated regions of interest;
- (i) screening compounds against the mutated regions of interest from step (h) and wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (j) identifying the compounds from step (i) that bind to the mutated regions of interest from step (h) and the wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (k) screening the compounds from step (j) against human 16S rRNA; and
- (l) identifying the drug candidates from step (k) that do not bind to the human 16S rRNA, thereby identifying drug candidates.

Another aspect of the present invention is a method for identifying drug candidates comprising:

- (a) transforming a first set of host cells with a first set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a first green fluorescent protein gene; wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, a mutant helix 9 sequence, and a first mutant Anti-Shine-Dalgarno sequence; and said first green fluorescent protein gene comprises a first mutant Shine-Dalgarno sequence; and wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; thereby forming a first set of transformed host cells;
- (b) isolating from the first set of transformed host cells those host cells which express the green fluorescent protein gene product;
- (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b) to identify regions of interest, wherein the regions of interest comprise sequences of one or more nucleic acids which are conserved in each first mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced;
- (d) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (d) are mutated; and each mutant *Pseudomonas aeruginosa* 16S rRNA gene further comprises a mutant helix 9 sequence, and a second mutant Anti-Shine-Dalgarno sequence;
- (e) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (d) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second green fluorescent protein having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;
- (f) transforming a second set of host cells with the plasmids from step (e), thereby forming a second set of transformed host cells;
- (g) isolating from the second set of transformed host cells from step (f) those host cells which express the green fluorescent protein gene product;
- (h) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (g), to identify the mutated regions of interest;

28

- (i) screening compounds against the mutated regions of interest from step (h) and wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (j) identifying the compounds from step (i) that bind to the mutated regions of interest from step (h) and the wildtype *Pseudomonas aeruginosa* 16S rRNA;
- (k) screening the compounds from step (j) against human 16S rRNA; and
- (l) identifying the drug candidates from step (k) that do not bind to the human 16S rRNA, thereby identifying drug candidates.

This invention is further illustrated by the following examples, which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and Appendices, are incorporated herein by reference.

## EXEMPLIFICATION

### EXAMPLE 1

#### Identification of Mutant SD AND Mutant ASD Combinations

It has been shown that by coordinately changing the SD and ASD, a particular mRNA containing an altered SD could be targeted to ribosomes containing the altered ASD. This and all other efforts to modify the ASD, however, have proved lethal, as cells containing these mutations died within two hours after the genes containing them were activated.

Using random mutagenesis and genetic selection, mutant SD-ASD combinations are screened in order to identify non-lethal SD-ASD combinations. The mutant SD-ASD mutually compatible pairs are set forth in FIGS. 10, 11 and 12. The mutually compatible pairs of mutant sequences interact as pairs in the form of RNA. The novel mutant SD-ASD sequence combinations of the present invention permit translation of only the mRNAs containing the altered SD sequence.

### EXAMPLE 2A

#### Construction of the Pa 16S pRNA228 Plasmid

A plasmid construct of the present invention identified as the Pa 16S pRNA228 plasmid, is set forth in FIGS. 1 and 6. *E. coli* cells contain a single chromosome with seven copies of the rRNA genes and all of the genes for the ribosomal proteins. The plasmid, Pa 16S pRNA228, in the cell contains a genetically engineered copy of one of the 16S rRNA genes from *P. aeruginosa* and two genetically engineered genes that are not normally found in *E. coli*, referred to herein as a "selectable markers." One gene encodes the protein chloramphenicol acetyltransferase (CAT). This protein renders cells resistant to chloramphenicol by chemically modifying the antibiotic. Another gene, the Green Fluorescent Protein (GFP), is also included in the system. GFP facilitates high-throughput functional analysis. The amount of green light produced upon irradiation with ultraviolet light is proportional to the amount of GFP present in the cell.

Ribosomes from Pa 16S pRNA228 have an altered ASD sequence. Therefore, the ribosomes can only translate mRNAs that have an altered SD sequence. Only two genes in the cell produce mRNAs with altered SD sequences that may be translated by the plasmid-encoded ribosomes: the

CAT and GFP gene. Mutations in rRNA affect the ability of the resulting mutant ribosome to make protein. The present invention thus provides a system whereby the mutations in the plasmid-encoded rRNA gene only affect the amount of GFP and CAT produced. A decrease in plasmid ribosome function makes the cell more sensitive to chloramphenicol and reduces the amount of green fluorescence of the cells. Translation of the other mRNAs in the cell is unaffected since these mRNAs are translated only by ribosomes that come from the chromosome. Hence, cells containing functional mutants may be identified and isolated via the selectable marker.

#### EXAMPLE 2B

##### Construction of the Pa 16S pRNA228 Plasmid

A plasmid construct of the present invention identified as the Pa 16S Ec H9 pRNA228 plasmid, is set forth in FIGS. 2 and 7. *E. coli* cells contain a single chromosome with seven copies of the rRNA genes and all of the genes for the ribosomal proteins. The plasmid, Pa 16S Ec H9 pRNA228, in the cell contains a genetically engineered copy of one of the 16S rRNA genes from *P. aeruginosa* wherein the *P. aeruginosa* helix 9 sequence has been replaced with the corresponding *E. coli* sequence, and two genetically engineered genes that are not normally found in *E. coli*, referred to herein as a "selectable markers." One gene encodes the protein chloramphenicol acetyltransferase (CAT). This protein renders cells resistant to chloramphenicol by chemically modifying the antibiotic. Another gene, the Green Fluorescent Protein (GFP), is also included in the system. GFP facilitates high-throughput functional analysis. The amount of green light produced upon irradiation with ultraviolet light is proportional to the amount of GFP present in the cell.

Ribosomes from Pa 16S Ec H9 pRNA228 have an altered ASD sequence. Therefore, the ribosomes can only translate mRNAs that have an altered SD sequence. Only two genes in the cell produce mRNAs with altered SD sequences that may be translated by the plasmid-encoded ribosomes: the CAT and GFP gene. Mutations in rRNA affect the ability of the resulting mutant ribosome to make protein. The present invention thus provides a system whereby the mutations in the plasmid-encoded rRNA gene only affect the amount of GFP and CAT produced. A decrease in plasmid ribosome function makes the cell more sensitive to chloramphenicol and reduces the amount of green fluorescence of the cells. Translation of the other mRNAs in the cell is unaffected since these mRNAs are translated only by ribosomes that come from the chromosome. Hence, cells containing functional mutants may be identified and isolated via the selectable marker.

#### EXAMPLE 2C

##### Construction of the pKan5-T1T2 and pKanPa-S20 Plasmid

A plasmid construct of the present invention identified as the pKan5-T1T2 plasmid, is set forth in FIGS. 3 and 8. The plasmid, pKan5-T1T2, is a PACYC177 derivative. It was used in the preparation of the pKanPa-S20 plasmid, which encodes the *P. aeruginosa* S20 protein, which is set forth in FIGS. 4 and 9.

#### EXAMPLE 3

##### Genetic System for Functional Analysis of Ribosomal RNA

Identification of Functionally Important Regions of rRNA. Functionally important regions of rRNA molecules that may be used as drug targets using a functional genomics approach may be identified through a series of steps. Namely, in step I.a., the entire rRNA gene is randomly mutated using error-prone PCR or another generalized mutational strategy. In step I.b., a host cell is then transformed with a mutagenized plasmid comprising: an rRNA gene having a mutant ASD sequence, at least one mutation in said rRNA gene, and a genetically engineered gene which encodes a selectable marker having a mutant SD sequence, and production of the rRNA genes from the plasmid are induced by growing the cells in the presence of IPTG. In step I.c., the CAT gene is used to select mutants that are functional by plating the transformed cells onto growth medium containing chloramphenicol. In step I.d., individual clones from each of the colonies obtained in step I.c. are isolated. In step I.e., the plasmid DNA from each of the individual clones from step I.d. is isolated. In step I.f., the rRNA genes contained in each of the plasmids that had been isolated in step I.e. are sequenced. In step I.g., the function of each of the mutants from step I.f. is assessed by measuring the amount of GFP produced by each clone from step I.e. upon induction of the rRNA operon. In step I.h., a functional genomics database is assembled containing the sequence and functional data from steps I.f. and I.g. In step I.i., functionally important regions of the rRNA gene that will serve as drug targets are identified. Functionally important regions may be identified by comparing the sequences of all of the functional genomics database constructed in step I.g. and correlating the sequence with the amount of GFP protein produced. Contiguous sequences of three or more rRNA nucleotides, in which substitution of the nucleotides in the region produces significant loss of function, will constitute a functionally important region and therefore a potential drug target. Helix 9 (SEQ ID NO 1 and SEQ ID NO 2) is a functional important region, as shown herein, and is therefore a potential drug target.

Isolation of Functional Variants of the Target Regions. A second aspect of the invention features identification of mutations of the target site that might lead to antibiotic resistance using a process termed, "instant evolution", as described below. In step II.a., for a given target region identified in step I.i., each of the nucleotides in the target region is simultaneously randomly mutated using standard methods of molecular mutagenesis, such as cassette mutagenesis or PCR mutagenesis, and cloned into the plasmid of step I.b. to form a plasmid pool containing random mutations at each of the nucleotide positions in the target region. In step II.b., the resulting pool of plasmids containing random mutations from step II.a. is used to transform *E. coli* cells and form a library of clones, each of which contains a unique combination of mutations in the target region. In step II.c., the library of mutant clones from step II.b. is grown in the presence of IPTG to induce production of the mutant rRNA genes. In step II.d., the induced mutants are plated on medium containing chloramphenicol, and CAT is used to select clones of rRNA mutants containing nucleotide combinations of the target region that produce functional ribosomes. In step II.e., the functional clones isolated in step II.d. are sequenced and GFP is used to measure ribosome

function in each one. In step II.f., the data from step II.e. are incorporated into a mutational database.

Isolation of Drug Leads. In step III.a., the database in step II.f. is analyzed to identify functionally-important nucleotides and nucleotide motifs within the target region. In step III.b., the information from step III.a. is used to synthesize a series of oligonucleotides that contain the functionally important nucleotides and nucleotide motifs identified in step III.a. In step III.c., the oligonucleotides from step III.b. are used to sequentially screen compounds and compound libraries to identify compounds that recognize (bind to) the functionally important sequences and motifs. In step III.d., compounds that bind to all of the oligonucleotides are counterscreened against oligonucleotides and/or other RNA containing molecules to identify drug candidates. "Drug candidates" are compounds that 1) bind to all of the oligonucleotides containing the functionally important nucleotides and nucleotide motifs, but do not bind to molecules that do not contain the functionally important nucleotides and nucleotide motifs and 2) do not recognize human ribosomes. Drug candidates selected by the methods of the present invention therefore recognize all of the functional variants of the target sequence, i.e., the target cannot be mutated in a way that the drug cannot bind, without causing loss of function to the ribosome.

#### EXAMPLE 4

##### Genetic System for Studying Protein Synthesis

All plasmids are maintained and expressed in *E. coli* DH5 (e.g. supE44, hsdR17, recA1, endA1, gyrA96, thi-1 and relA1; Hanahan, D. (1983) *J. Mol. Biol.* 166:557-580). To induce synthesis of plasmid-derived rRNA from the lacUV5 promoter, IPTG is added to a final concentration of 1 mM. Chloramphenicol acetyltransferase activity will be determined essentially as described by Nielsen et al. (1989) *Anal. Biochem.* 179: 19-23. Cultures for CAT assays are grown in LB-Ap100. MIC will be determined by standard methods in microtiter plates as described in Lee, K., et al. (1997) *J. Mol. Biol.* 269: 732-743. Procedures are followed as in outlined in Example 4 of Cunningham et al. (WO 2004/003511).

#### EXAMPLE 5

##### In Vivo Determination of RNA Structure-Function Relationships

Bacterial strains and media. Plasmids are maintained and expressed in *E. coli* DH5 (e.g. supE44, hsdR17, recA1, endA1, gyrA9 and thi-1; Hanahan, D. (1983) *J. Mol. Biol.* 166:557-580). Cultures are grown in LB medium (Luria, S. E. & Burrous, J. W. (1957) *J. Bacteriol.* 74:461-476) or LB medium containing 100 µg/ml ampicillin (LB-Ap100). To induce synthesis of plasmid-derived rRNA from the lacUV5 promoter, IPTG is added to a final concentration of 1 mM at the times indicated in each experiment. Strains are transformed by electroporation (Dower, W. J., et al. (1988) *Nucl. Acids Res.* 16: 6127) using a Gibco-BRL Cell Porator. Transformants are grown in SOC medium (Hanahan, 1983,

supra) for one hour prior to plating on selective medium to allow expression of plasmid-derived genes.

Chloramphenicol acetyltransferase assays. CAT activity is determined essentially as described (Nielsen, D. A. et al. (1989) *Anal. Biochem.* 60:191-227). Cultures for CAT assays will be grown in LB-Ap100. Briefly, 0.5 ml aliquots of mid-log cultures (unless otherwise indicated) is added to an equal volume of 500 mM Tris-HCl (pH8) and lysed using 0.01% (w/v) SDS and chloroform (Miller, J. H. (1992) *A Short Course in Bacterial Genetics*, (Miller, J. H., ed.), pp. 71-80, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). The resulting lysate is either be used directly or diluted in assay buffer prior to use. Assay mixtures contain cell extract (5 µl or 10 µl), 250 mM Tris (pH 8), 214 µM butyryl-coenzyme A (Bu-CoA), and 40 µM [<sup>3</sup>H]chloramphenicol in a 125 µl volume. Two concentrations of lysate are assayed for one hour at 37° C. to ensure that the signal was proportional to protein concentrations. The product, butyryl-<sup>3</sup>H]chloramphenicol is extracted into 2,6,10,14-tetramethylpentadecane:xylene (2:1) and measured directly in a Beckman LS-3801 liquid scintillation counter. Blanks are prepared exactly as described above, except that uninoculated LB medium was used instead of culture.

Minimum inhibitory concentration determination. MICs are determined by standard methods in microtiter plates or on solid medium. Overnight cultures grown in LB-Ap100 are diluted and induced in the same medium containing 1 mM IPTG for three hours. Approximately 10<sup>4</sup> induced cells are then added to wells (or spotted onto solid medium) containing LB-Ap100 +IPTG (1 mM) and chloramphenicol at increasing concentrations. Cultures are grown for 24 hours and the lowest concentration of chloramphenicol that completely inhibited growth is designated as the MIC.

Oligoribonucleotide synthesis. Oligoribonucleotides are synthesized on solid support with the phosphoramidite method (Capaldi, D. & Reese, C. (1994) *Nucl. Acids Res.* 22:2209-2216) on a Cruachem PS 250 DNA/RNA synthesizer. Oligomers are removed from solid support and deprotected by treatment with ammonia and acid following the manufacturer's recommendations. The RNA is purified on a silica gel Si500F TLC plate (Baker) eluted for five hours with n-propanol/ammonia/water (55:35:10, by vol.). Bands are visualized with an ultraviolet lamp and the least mobile band was cut out and eluted three times with 1 ml of purified water. Oligomers are further purified with a Sep-pak C-18 cartridge (Waters) and desalted by continuous-flow dialysis (BRL). Purities are checked by analytical C-8 HPLC (Perceptive Biosystems).

#### INCORPORATION BY REFERENCE

All of the references including, without limitation, U.S. patents, U.S. patent application publications, published international applications and journal articles cited herein are hereby incorporated by reference.

#### EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 155

<210> SEQ ID NO 1

<211> LENGTH: 19

<212> TYPE: RNA

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 1

auaacgucgc aagaccaaa 19

<210> SEQ ID NO 2

<211> LENGTH: 19

<212> TYPE: RNA

<213> ORGANISM: Pseudomonas aeruginosa

<400> SEQUENCE: 2

auacguccug agggagaaa 19

<210> SEQ ID NO 3

<211> LENGTH: 13115

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic nucleotide sequence

<400> SEQUENCE: 3

ttacgccccg ccctgccact catcgcagta ctgttgaat tcattaagca ttctgccgac 60

atggaagcca tcacagacgg catgatgaac ctgaatcgcc ageggcatca gcaccttgtc 120

gccttgcgta taatatttgc ccatggtgaa aacgggggcg aagaagttgt ccatattggc 180

cacgtttaa tcaaaactgg tgaaactcac ccagggattg gctgagacga aaaacatatt 240

ctcaataaac cctttaggga aataggccag gttttcaccg taacacgcca catcttgcca 300

atatatgtgt agaaactgcc ggaaatcgtc gtggtattca ctccagagcg atgaaaacgt 360

ttcagtttgc tcatgaaaaa cgggtgaaca agggtaaca ctatccata tcaccagctc 420

accgtctttc attgccatac ggaattccgg atgagcattc atcaggcggg caagaatgtg 480

aataaaggcc ggataaaact tgtgcttatt tttctttacg gtctttaaaa aggccgtaat 540

atccagctga acggtctggt tataggtaca ttgagcaact gactgaaatg cctcaaaatg 600

ttctttacga tgccattggg atatatcaac ggtggtatat ccagtgattt ttttctccat 660

ttgcgagggg atatgaaagc ggccgcttcc acacattaaa ctagttcgat gattaattgt 720

caacagctcg cggcgggcac ctgcgtaacg gattcaccac tccaagaatt ggagccaatc 780

gattcttgcg gagaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgctg 840

ccgccatctc cagcagccgc acgcgggcga tctcgggcag cgttgggtcc tggccacggg 900

tgcgcatgat cgtgctcctg tcggtgagga cccggctagg ctggcggggg tgecttactg 960

gtagcagaa tgaatcaccg atacgcgagc gaacgtgaag cgactgctgc tgcaaaacgt 1020

ctgcgacctg agcaacaaca tgaatggtct tcggtttccg tgtttcgtaa agtctggaaa 1080

cgcggaagtc agcgccctgc accattatgt tccggatctg ggtaccgca ttcacagttc 1140

tccgcaagaa tcgattggtc ccaattcttg gagtggtgaa tccgtagcg aggtgcccgc 1200

ggcgagctgt tgacaattaa tcatcgaact agtttaatgt gtggaagcgg ccgctttcat 1260

atccctccgc aatggagaa aaaaactact gctagcaaaag gagaagaact tttcactgga 1320

gttgtcccaa ttcttgttga attagatggt gatgttaatg ggcacaaatt ttctgtcagt 1380

-continued

---

ggagaggggtg	aaggtgatgc	tacatacggg	aagcttacc	ttaaatttat	ttgcactact	1440
ggaaaactac	ctgttccatg	gccaacactt	gtcactactt	tctcttatgg	tgttcaatgc	1500
ttttcccgtt	atccggatca	tatgaaacgg	catgactttt	tcaagagtgc	catgccccgaa	1560
ggttatgtac	aggaacgcac	tatatctttc	aaagatgacg	ggaactacaa	gacgcgtgct	1620
gaagtcaagt	ttgaaggtga	tacccttggt	aatcgtatcg	agttaaaagg	tattgatttt	1680
aaagaagatg	gaaacattct	cggacacaaa	ctcgagtaca	actataactc	acacaatgta	1740
tacatcacgg	cagacaaaac	aaagaatgga	atcaaagcta	acttcaaaat	tcgccacaac	1800
atgaagatg	gatccgttca	actagcagac	cattatcaac	aaaatactcc	aattggcgat	1860
ggccctgtcc	ttttaccaga	caaccattac	ctgtcgacac	aatctgcctt	ttcgaaagat	1920
cccaacgaaa	agcgtgacca	catggctcct	cttgagtttg	taactgctgc	tgggattaca	1980
catggcatgg	atgagctcta	caaataatct	agtcgtagcg	ccgatggtag	tgtggggctt	2040
cccctatgca	gagtagggaa	ctgccaggca	tcaaataaaa	cgaaaggctc	agtcgaaaga	2100
ctgggccttt	cgttttatct	gttgtttgtc	ggtgaacgct	ctcctgagta	ggacaaatcc	2160
gccgggagcg	gatttgaacg	ttgcgaagca	acggcccggg	gggtggcggg	caggacgccc	2220
gccataaact	gccaggcatc	aaattaagca	gaaggccatc	ctgacggatg	gcctttttgc	2280
gtttctacaa	actcttctcg	tcgtcagtgc	aggtaccgag	ctcgaattca	ctggccgctc	2340
ttttacaacg	tcgtgactgg	gaaaaccctg	gcgttaccga	acttaatcgc	cttgcagcac	2400
atcccccttt	cgccaggcat	cgcaggatgc	tgetggctac	cctgtggaac	acctacatct	2460
gtattaacga	agcgtggca	ttgaccctga	gtgatttttc	tctggtccc	ccgcatccat	2520
accgccagtt	gtttaccctc	acaacgttcc	agtaaccggg	catgttcac	atcagtaacc	2580
cgtatcgtga	gcctcctctc	tcgtttcctc	ggtatcatta	ccccatgaa	cagaaattcc	2640
cccttacacg	gaggcatcaa	gtgaccaaac	aggaaaaaac	cgccctaac	atggcccgtt	2700
ttatcagaag	ccagacatta	acgcttctgg	agaaactcaa	cgagctggac	gcggatgaac	2760
aggcagacat	ctgtgaatcg	cttcacgacc	acgctgatga	gctttaccgc	agctgcctcg	2820
cgcgtttcgg	tgatgacggg	gaaaacctct	gacacatgca	gctcccggag	acggtcacag	2880
cttgtctgta	agcggatgcc	gggagcagac	aagcccgtca	gggcgcgtca	gcgggtggtg	2940
gcgggtgctg	gggcgcagcc	atgaccagct	cacgtagcga	tagcggagtg	tatactggct	3000
taactatgcg	gcatcagagc	agattgtact	gagagtgcac	catatgcggt	gtgaaatacc	3060
gcacagatgc	gtaaggagaa	aataccgcat	caggcgtctt	tccgcttctt	cgctcactga	3120
ctcgtcgcgc	tcggtcgttc	ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	3180
acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	3240
aaaggccagg	aaccgtaaaa	aggccgcggt	gctggcggtt	ttccataggc	tcggcccccc	3300
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgca	caggactata	3360
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	3420
gcttacggga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	3480
acgctgtagg	tatctcagtt	cgggtgtaggt	cgctcgtctc	aagctgggct	gtgtgcacga	3540
acccccggtt	cagcccagcc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaaccc	3600
ggtaagacac	gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	3660
gtatgtaggc	ggtgctacag	agttcttgaa	gtggggcctt	aactacggct	acactagaag	3720
gacagtatct	ggtatctgcg	ctctgctgaa	gccagttacc	ttcgaaaaaa	gagttggtag	3780

-continued

---

ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcggtggt	ttttttgttt	gcaagcagca	3840
gattacggc	agaaaaaag	gatctcaaga	agatcctttg	atctttteta	cggggtctga	3900
cgctcagtgg	aacgaaaact	cacgttaagg	gattttggtc	atgagattat	caaaaaggat	3960
cttcacctag	atccttttaa	attaaaaatg	aagttttaaa	tcaatctaaa	gtatatatga	4020
gtaaacttgg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	cagcgatctg	4080
tctatttcgt	tcatccatag	ttgcctgact	ccccgctgtg	tagataacta	cgatacggga	4140
gggcttacca	tctggcccca	gtgctgcaat	gataccgcca	gacccacgct	caccggctcc	4200
agatttatca	gcaataaacc	agccagccgg	aaggccggag	cgcagaagtg	gtcctgcaac	4260
tttatccgcc	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	gtagttcgcc	4320
agttaatagt	ttgcgcaacg	ttgttgccat	tgctgcaggc	atcgtggtgt	cacgctcgtc	4380
gtttggtag	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	catgatcccc	4440
catgttgtgc	aaaaaagcgg	ttagctcctt	cggtcctccg	atcgttgtca	gaagtaagtt	4500
ggccgcagtg	ttatcaactca	tggttatggc	agcaactgcat	aattctctta	ctgtcatgcc	4560
atccgtaaga	tgcttttctg	tgactggtga	gtactcaacc	aagtcattct	gagaatagtg	4620
tatgcccga	cagagttgct	cttgcccggc	gtcaaacagg	gataataaccg	cgccacatag	4680
cagaacttta	aaagtgctca	tcattggaaa	acgttcttcg	gggcgaaaac	tctcaaggat	4740
cttaccgctg	ttgagatcca	gttcgatgta	accactcgt	gcacccaact	gatcttcagc	4800
atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaca	ggaaggcaaa	atgccgcaaa	4860
aaaggaata	aggcgacac	ggaaatggtg	aatactcata	ctcttccttt	ttcaatatta	4920
ttgaagcatt	tatcaggggt	attgtctcat	gagcggatac	atatttgaat	gtatttagaa	4980
aaataaacia	ataggggttc	cgcgcacatt	tccccgaaaa	gtgccacctg	acgtctaaga	5040
aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggc	cctttcgtct	5100
tcaagaattc	tcatgtttga	cagcttatca	tcgataagct	ttaatgcggt	agtttatcac	5160
agttaaattg	ctaacgcagt	caggcacctg	gtatgaaatc	taacaatgcg	ctcatcgta	5220
tctcggcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	gtactgccgg	5280
gcctcttgcg	ggatatcgtc	cattccgaca	gcategccag	tcactatggc	gtgctgctag	5340
cgctatatgc	gttgatgcaa	tttctatgcg	caccgcttct	cggagcactg	tccgaccgct	5400
ttggccgccg	cccagtcctg	ctcgcttcgc	tacttgagac	cactatcgac	tacgcgatca	5460
tgccgaccac	accgctcctg	tggatcccag	acgagttaag	tcaccatacg	ttagtacagg	5520
ttgccactct	tttggcagac	gcagacctac	ggctacaata	gcgaagcggc	cctgggtattc	5580
atgtttaaaa	atactgtcgc	gatagccaaa	acggcactct	ttggcagtta	agcgcacttg	5640
cttgccctgtc	gccagttcaa	cagaatcaac	ataagcgcaa	actcgctgta	attctacgcc	5700
ataagcacca	atattctgga	taggtgatga	gccgacacaa	ccaggaatta	atgccagatt	5760
ttccagacca	ggcatacctt	cctgcaaagt	gtattttacc	agacgatgcc	agttttctcc	5820
ggctcctaca	tgtaataacc	acgcatcagg	ttcatcatga	atttcgatac	ctttgatccg	5880
gttgatgatc	cctgcaggcc	cttaaggcca	tttaaatggc	gcgccgatca	atgccaaatg	5940
tgttccaggg	ttttaaggag	tggttcatag	ctgctttcct	gatgcaaaaa	cgaggctagt	6000
ttaccgtatc	tgtgggggga	tggctttag	atatgacgac	aggaagagtt	tgtagaaacg	6060
caaaaaggcc	atccgtcagg	atggccttct	gcttaatttg	atgcctggca	gtttatggcg	6120

-continued

---

ggcgtcctgc	ccgccaccct	ccgggcccgtt	gcttcgcaac	gttcaaatcc	gctcccggcg	6180
gatttgctct	actcaggaga	gcgttcaccg	acaaacaaca	gataaaacga	aaggcccagt	6240
ctttcgactg	agcctttcgt	tttatttgat	gcctggcagt	tcctactct	cgcatgggga	6300
gacccacac	taccatcggc	gctacggcgt	ttcacttctg	agttcggcat	ggggtcaggt	6360
gggaccaccg	cgctactgcc	gccaggcaaa	ttctgtttta	tcagaccgct	tctgcgttct	6420
gatttaactc	gtatcaggct	gaaaatcttc	tctcatccgc	caaacagct	tcggcgttgt	6480
aaggtttaagc	ctcacggttc	attagtaccg	gttagctcaa	cgcatcgctg	cgcttacaca	6540
cccggcctat	caacgtcgtc	gtcttcaacg	ttccttcagg	acccttaaag	ggtcagggag	6600
aactcatctc	ggggcaagtt	tcgtgcttag	atgctttcag	cacttatctc	ttccgcattt	6660
agctaccggg	cagtgcatt	ggcatgacaa	cccgaacacc	agtgatgcgt	ccactccggt	6720
cctctcgtac	taggagcagc	ccccctcagt	tctccagcgc	ccacggcaga	tagggaccga	6780
actgtctcac	gacgttctaa	accagctcgc	cgtaccactt	taaattggcg	acagccatac	6840
ccttgggacc	tacttcagcc	ccaggatgtg	atgagccgac	atcgaggtgc	caaacaccgc	6900
cgtcgatatg	aactcttggg	cggtatcagc	ctgttatccc	cggagtacct	tttatccggt	6960
gagcgatggc	ccttccattc	agaaccaccg	gatcactatg	acctgcttcc	gcacctgctc	7020
gcgcctgcac	gctcgcagtc	aagctggcct	atgccattgc	actaacctcc	tgatgtccga	7080
ccaggattag	ccaaccttcg	tgtcctcccg	ttactcttta	ggaggagacc	gccccagtca	7140
aactaccac	cagacactgt	ccgcaaccgc	gattacgggt	caacgttaga	acatcaaaca	7200
ttaaagggtg	gtatttcaag	gtcggctcca	tgagactgg	cgtccacact	tcaaagcctc	7260
ccacctatcc	tacacatcaa	ggctcaatgt	tcagtgtcaa	gctatagtaa	aggttcacgg	7320
ggtctttccg	tcttgccgcg	ggtacactgc	atcttcacag	cgagttcaat	ttcactgagt	7380
ctcgggtgga	gacagcctgg	ccatcattac	gccattcgtg	caggtcggaa	cttaccgcac	7440
aaggaatttc	gctaccttag	gaccgttata	gttacggccg	ccgtttaccg	gggcttcgat	7500
caagagcttc	gcttgcgcta	accccatcaa	ttaaccttcc	ggcaccgggc	aggcgtcaca	7560
ccgtatacgt	ccactttcgt	gtttgcacag	tgctgtgttt	ttaataaaca	gttgcagcca	7620
gctggtatct	tcgactgatt	tcagctccat	ccgcgagggg	cctcacctac	atatcagcgt	7680
gccttctccc	gaagttacgg	caccatthttg	cctagttcct	tcacccgagt	tctctcaagc	7740
gccttgggat	tctctacctg	accacctgtg	tcggtttggg	gtacgatttg	atggtacctg	7800
atgcttagag	gcttttctctg	gaagcagggc	atgtgttget	tcagcaccgt	agtgcctcgt	7860
catcacgcct	cagccttgat	tttccggatt	tgcttgaaa	accagcctac	acgcttaaac	7920
ccggacaacc	gtcgcgccgc	caacatagcc	ttctccgtcc	ccccttcgca	gtaaacacaa	7980
gtacaggaat	attaacctgt	ttcccatcga	ctacgccttt	cggcctcgcc	ttaggggtcg	8040
actcacctg	ccccgattaa	cgttggacag	gaaccttgg	tcttccggcg	agcgggcttt	8100
tcacccgctt	tategttact	tatgtcagca	ttcgcacttc	tgatacctcc	agcatgcctc	8160
acagcacacc	ttcgcaggct	tacagaacgc	ttccctaccc	aacaacgcat	aagcgtcgtc	8220
gccgcagctt	ccgtgcatgg	tttagccccg	ttacatcttc	ccgcgaggcc	gactcgacca	8280
gtgagctatt	acgctttctt	taaattgatg	ctgcttctaa	gccaacatcc	tggtgtctg	8340
ggccttccca	catgthttcc	cacttaacca	tgactttggg	accttagctg	gcggctctggg	8400
ttgtttccct	cttcacgacg	gacgttagca	cccgcctgtg	gtctcccgctg	ataacattct	8460
ccggtattcg	cagtttgcac	ccggttggta	agtcgggatg	accccttgc	cgaaacagtg	8520

-continued

---

ctctaccccc	ggagatgaat	tcacgaggcg	ctacctaaat	agctttcggg	gagaaccagc	8580
tatctcccgg	tttgattggc	ctttcacccc	cagccacaag	tcatecgeta	atTTTTcaac	8640
attagtcggg	tcggctctcc	agttagtggt	acccaacctt	caacctgccc	atggctagat	8700
caccggggtt	cgggtctata	ccctgcaact	taacgcccag	ttaagactcg	gtttcccttc	8760
ggctccccta	ttcggttaac	cttgctacag	aatataagtc	gctgacctat	tatacaaaag	8820
gtacgcagtc	acacgcctaa	gcgtgctccc	actgcttgta	cgtacacggt	ttcaggttct	8880
ttttcactcc	cctcgcgggg	gttcttttcg	cctttccctc	acggctactg	ttcactatcg	8940
gtcagtcagg	agtatttagc	cttgagggat	ggcccccca	tattcagaca	ggataccacg	9000
tgtcccggcc	tactcatcga	gctcacagca	tgtgcatttt	tgtgtacggg	gctgtcacc	9060
tgtatcgcgc	gcctttccag	acgcttccac	taacacacac	actgattcag	gctctgggct	9120
gctcccgggt	cgctcgcgc	tactggggga	atctcgggtg	atctcttttc	ctcgggggtac	9180
ttagatggtt	cagttcccc	ggttcgcctc	attaacctat	ggattcagtt	aatgatagtg	9240
tgtcgaaaaca	cactggggtt	ccccattcgg	aaatcgcggg	ttataacggt	tcatatcacc	9300
ttaccgacgc	ttatcgcaga	ttagcacgtc	cttcatcgcc	tctgactgcc	agggcatcca	9360
ccgtgtacgc	ttagtcgctt	aacctcacia	cccgaagatg	ttcttttcga	ttcatcatcg	9420
tgttgcgaaa	atTTGagaga	ctcacgaaca	actctcgttg	ttcagtggtt	caattttcag	9480
cttgatccag	atTTTTaaag	agcaaatatc	tcaaacatca	cccgaagatg	agTTTTgaga	9540
tattaaggtc	ggcgactttc	actcaciaaac	cagcaagtgg	cgtcccctag	gggattcгаа	9600
cccctgttac	cgccgtgaaa	gggcgggtgc	ctgggcctct	agacgaaggg	gacacgaaaa	9660
ttgcttatca	cgcgttgctg	gatatttttcg	tgtagggtga	gctttcatta	atagaaagcg	9720
aacggcctta	ttctcttcag	cctcactccc	aacgcgtaaa	cgccttgctt	ttcactttct	9780
atcagacaat	ctgtgtgagc	actacaaagt	acgcttcttt	aaggtaatcc	catgatccaa	9840
ccgcaggttc	ccctacgggt	accttggttac	gacttcaccc	cagtcatgaa	tcactccgtg	9900
gtaaccgtcc	cccttgcggt	tagactagct	acttctggag	caaccactc	ccatgggtgtg	9960
acgggcgggtg	tgtacaaggc	ccgggaacgt	attcacctgtg	acattctgat	tcacgattac	10020
tagcgattcc	gacttcacgc	agtcgagttg	cagactgcga	tccggactac	gatcgggttt	10080
atgggattag	ctccacctcg	cggcttgga	accctttgta	ccgaccattg	tagcacgtgt	10140
gtagccctgg	ccgtaagggc	catgatgact	tgacgtcacc	cccaccttc	tccggtttgt	10200
caccggcagt	ctccttagag	tgccaccccg	aggtgctggg	aactaaggac	aagggttgcg	10260
ctcgttacgg	gacttaacc	aacatctcac	gacacgagct	gacgacagcc	atgcagcacc	10320
tgtgtctgag	ttcccgaagg	caccaatcca	tctctggaaa	gttctcagca	tgtcaaggcc	10380
aggtaaagggt	cttcgcgttg	cttcgaatta	aaccacatgc	tccaccgctt	gtgcggggccc	10440
ccgtcaatcc	atTTGagttt	taaccttgcg	gccgtactcc	ccaggcggtc	gacttatcgc	10500
gttagctgcg	ccactaagat	ctcaaggatc	ccaacggcta	gtcgacatcg	tttacggcgt	10560
ggactaccag	ggatctaat	cctgtttgct	ccccacgctt	tcgcacctca	gtgtcagtat	10620
cagtccagggt	ggcgccttc	gccactgggtg	ttccttccca	tatctacgca	tttcaccgct	10680
acacaggaaa	ttccaccacc	ctctaccgta	ctctagctca	gtagttttgg	atgcagttcc	10740
caggttgagc	ccggggattt	cacatccaac	ttgctgaacc	acctaegcgc	gctttacgcc	10800
cagtaattcc	gattaacgct	tgcacccttc	gtattaccgc	ggctgctggc	acgaagttag	10860



-continued

---

```

ccggtgctta ttctgttggc aacgtcaaaa cagcaaggta ttaacttact gcccttcctc 10920
ccaacttaaa gtgctttaca atccgaagac cttcttcaca cacgcggcat ggctggatca 10980
ggctttcgcc cattgtccaa tattccccac tgctgcctcc cgtaggagtc tggaccgtgt 11040
ctcagttcca gtgtgactga tcctcctctc agaccagtta cggatcgtcg ccttggtagg 11100
cctttacccc accaactagc taatccgacc taggctcctc tgatagcgtg aggtccgaag 11160
atccccact ttctcctca ggacgtatgc ggtattagcg cccgtttccg gacgttatcc 11220
cccactacca ggcagattcc taggcattac tcaccctgac gccgctgaat ccaggagcaa 11280
gctcccttca tccgctcgac ttgcatgtgt taggcctgcc gccagcgttc aatctgagcc 11340
atgatcaaac tcttcaattt aaaagtttga cgctcaaaga attaaacttc gtaatgaatt 11400
acgtgttcac tcttgagact tggattcat tttctgctt gcgacgttaa gaatccgtat 11460
cttcgagtgc ccacacagat tgtctgataa attgttaaag agcagtgccg cttegtttt 11520
tctcagcggc cgctgtgtga aattgttctc cgctcacaat tccacacatt atacgagccg 11580
gaagcataaa gtgtaaagcc tgggggtgct aatgagttag ctaactcaca ttaattgcgt 11640
tgcgctcact gcccgtttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg 11700
gccaacgcgc ggggagagcc ggtttgcgta ttgggcgcca ggggtggttt tcttttcacc 11760
agtgagacgg gcaacagctg attgcccttc accgcctggc cctgagagag ttgcagcaag 11820
cggccacgc tggtttgccc cagcaggcga aaatcctgtt tgatggtggt tgacggcggg 11880
atataacatg agctgtcttc ggtatcgtcg tatcccacta ccgagatctc cgcaccaacg 11940
cgcagcccgg actcggtaat ggcgcgcatt gcgcccagcg ccatctgatc gttggcaacc 12000
agcatcgcag tgggaacgat gccctcattc agcatttgca tggtttgtt aaaaccggac 12060
atggcactcc agtcgccttc cgttccgct atcggctgaa tttgattgcg agtgagatat 12120
ttatgccagc cagccagacg cagacgcgcc gagacagaac ttaatgggcc cgctaacagc 12180
gcgatttgct ggtgbcccaa tgcgaccaga tgctccacgc ccagtcgctg accgtcttca 12240
tgggagaaaa taatactgtt gatgggtgct tggtcagaga catcaagaaa taacgccgga 12300
acattagtgc aggcagcttc cacagcaatg gcatcctggt catccagcgg atagttaata 12360
atcagcccac tgaccggtg cgcgagaaga ttgtgcaccg ccgctttaca ggcttcgacg 12420
ccgcttcggt ctaccatcga caccaccacg ctggcaccga gttgatcggc gcgagattta 12480
atcgcgcgga caatttgca cggcgcgtgc agggccagac tggaggtggc aacgccaatc 12540
agcaacgact gtttgcccgc cagttgttgt gccacgcggt tgggaatgta attcagctcc 12600
gccatcgcg cttccacttt ttcccgcgtt ttcgcagaaa cgtggctggc ctggttcacc 12660
acgcgggaaa cggctctgata agagacaccg gcatactctg cgacatcgtg taacgttact 12720
ggtttcacat tcaccacct gaattgactc tcttccgggc gctatcatgc cataccgca 12780
aaggttttgc accattcgat ggtgtcggat cctctagccg gaccacttg cggccacgat 12840
ccgtccgccc taaggetcat accgttaatt attccccccc acgggagacc tgagcaaact 12900
gccctcagcc atttgagaag cacagggtca cactgcttcg ggtagtcaat aaaccggtaa 12960
accagcaata gacataagcg gctatttaac gaccctgcc tgaaccgacg accgggtcga 13020
atgtgcttcc gaatttctgc cattcatccg cttattatca cttattcagg cgtagcacca 13080
ggcgtttaag ggcaccaata actgccttaa aaaaa 13115

```

&lt;210&gt; SEQ ID NO 4

&lt;211&gt; LENGTH: 13115

-continued

---

```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      nucleotide sequence

<400> SEQUENCE: 4
ttacgccccg cctgcccact catcgcagta ctgttgtaat tcattaagca ttctgcccac    60
atggaagcca tcacagacgg catgatgaac ctgaatcgcc agcggcatca gcaccttgtc    120
gccttgcgta taatatttgc ccatggtgaa aacggggggcg aagaagttgt ccatattggc    180
cacgtttaa tcaaaactgg tgaaactcac ccagggattg gctgagacga aaaacatatt    240
ctcaataaac cctttaggga aataggccag gttttcaccg taacacgcca catcttgcca    300
atatatgtgt agaaactgcc ggaaatcgtc gtggtattca ctccagagcg atgaaaacgt    360
ttcagtttgc tcatggaaaa cgggtgaaca aggggtgaaca ctatcccata tcaccagctc    420
accgtctttc attgccatac ggaattccgg atgagcattc atcaggcggg caagaatgtg    480
aataaaggcc ggataaaact tgtgcttatt tttctttacg gtctttaaag aggccgtaat    540
atccagctga acggtctggt tataggtaca ttgagcaact gactgaaatg cctcaaaatg    600
ttctttacga tgccattggg atatatcaac ggtggtatat ccagtgattt ttttctccat    660
ttgcccgggg atatgaaagc ggccgcttcc acacattaaa ctagttcgat gattaattgt    720
caacagctcg ccggcggcac ctgcgtaacg gattcaccac tccaagaatt ggagccaatc    780
gattcttgcg gagaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgcgt    840
ccgccatctc cagcagccgc acgcccgcga tctcgggcag cgttggttcc tggccacggg    900
tgcgcatgat cgtgctcctg tcggtgagga cccggctagg ctggcggggg tgccctactg    960
gtagcagaa tgaatcaccg atacgcgagc gaacgtgaag cgactgctgc tgcaaaacgt   1020
ctgcgacctg agcaacaaca tgaatggtct tcggtttccg tgtttcgtaa agtctggaaa   1080
cgccgaagtc agcgccttgc accattatgt tccggatctg ggtaccgca ttcacagttc   1140
tccgcaagaa tcgattggct ccaattcttg gagtgtgaa tccgtagcg aggtgccgcc   1200
ggcgagctgt tgacaattaa tcatcgaact agtttaatgt gtggaagcgg ccgctttcat   1260
atccctccgc aatggagaa aaaaatcact gctagcaaag gagaagaact tttcactgga   1320
gttgccccaa ttcttggtga attagatggt gatgttaatg ggcacaaatt ttctgtcagt   1380
ggagaggggt aaggtgatgc tacatacggg aagcttacc ttaaatttat ttgcactact   1440
ggaaaactac ctgttccatg gccaacactt gtcaactact tctcttatgg tgttcaatgc   1500
ttttccggtt atccgatca tatgaaacgg catgactttt tcaagagtgc catgcccga   1560
ggttatgtac aggaacgcac tatatctttc aaagatgacg ggaactaca gacgcgtgct   1620
gaagtcaagt ttgaaggtga tacccttggt aatcgtatcg agttaaagg tattgatttt   1680
aaagaagatg gaaacattct cggacacaaa ctcgagtaca actataactc acacaatgta   1740
tacatcacgg cagacaaaca aaagaatgga atcaaagcta acttcaaaat tcgccacaac   1800
attgaagatg gatccgttca actagcagac cattatcaac aaaatactcc aattggcgat   1860
ggcctgtcc ttttaccaga caaccattac ctgtcgacac aatctgccct ttcgaaagat   1920
cccaacgaaa agcgtgacca catggctcct cttgagtttg taactgctgc tgggattaca   1980
catggcatgg atgagctcta caaataatct agtcgtagcg ccgatgtag tgtggggctc   2040
ccccatgcca gagtagggaa ctgccaggca tcaataaaaa cgaaaggctc agtcgaaaga   2100
ctgggccttt cgttttatct gttgtttgc ggtgaacgct ctctgagta ggacaaatcc   2160

```

-continued

---

gccgggagcg gatttgaacg ttgcgaagca acggcccgga ggggtggcggg caggacgccc	2220
gccataaact gccaggcatc aaattaagca gaaggccatc ctgacggatg gcctttttgc	2280
gtttctacaa actcttctcg tcgtcagtgc aggtaccgag ctccaattca ctggccgctcg	2340
ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac	2400
atcccccttt cgccaggcat cgcaggatgc tgctggctac cctgtggaac acctacatct	2460
gtattaacga agcgcctggca ttgaccctga gtgatttttc tctggctccc cgcacatccat	2520
accgccagtt gtttaccctc acaacgttcc agtaaccggg catgttcac c atcagtaacc	2580
cgtatcgtga gcacccctc tcgtttcctc ggtatcatta ccccatgaa cagaaattcc	2640
cccttacacg gaggcacaa gtgaccaaac aggaaaaaac cgccctaac atggcccgt	2700
ttatcagaag ccagacatta acgcttctgg agaaactcaa cgagctggac gcggatgaac	2760
aggcagacat ctgtgaatcg cttcacgacc acgctgatga gctttaccgc agctgcctcg	2820
cgcgtttcgg tgatgacggt gaaaacctct gacacatgca gctcccggag acggtcacag	2880
cttgtctgta agcggatgcc gggagcagac aagcccgtca gggcgcgtca gcgggtggtg	2940
gcgggtgctg gggcgcagcc atgaccagc cacgtagcga tagcggagtg tatactggct	3000
taactatgcg gcacagagc agattgtact gagagtgcac catatgcggt gtgaaatacc	3060
gcacagatgc gtaaggagaa aataccgcat caggcgtctc tccgcttctc cgctcactga	3120
ctcgtcgcgc tcggctcgtc ggctgcggcg agcggatca gctcactcaa aggcggtaat	3180
acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca	3240
aaaggccagg aaccgtaaaa aggcgcgctt gctggcgttt ttccataggc tccgcccc	3300
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccgca caggactata	3360
aagataccag gcgtttcccc ctggaagctc cctcgtcgc tctcctgttc cgaccctgcc	3420
gcttacggga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc	3480
acgctgtagg tatctcagtt cgggtgtagg ctctcgtcc aagctgggct gtgtgcacga	3540
acccccgct cagcccagc gctgcgcctt atccggtaac tategtcttg agtccaacc	3600
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	3660
gtatgtaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct aactagaag	3720
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag	3780
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca	3840
gattacgcgc agaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	3900
cgctcagtgg aacgaaaact cacgttaagg gatthttggtc atgagattat caaaaaggat	3960
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga	4020
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	4080
tctatttcgt tcatccatag ttgcctgact ccccgctcgt tagataacta cgatacggga	4140
gggcttacca tctggcccca gtgctgcaat gataccgcca gaccacgct caccggctcc	4200
agatttatca gcaataaacc agccagccgg aaggcccgag cgcagaagtg gtccctgcaac	4260
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc	4320
agttaatagt ttgcgcaacg ttggtgcat tgctgcaggc atcgtgggtg cacgctcgtc	4380
gtttggtatg gcttcattca gctccggctt ccaacgatca aggcgagtta catgatcccc	4440
catggtgtgc aaaaaagcgg ttagctcctt cggctcctcg atcgttgtca gaagtaagtt	4500

-continued

---

ggccgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	ctgtcatgcc	4560
atccgtaaga	tgcttttctg	tgactgggta	gtactcaacc	aagtcattct	gagaatagtg	4620
tatgcccgga	ccgagttgct	cttgcccggc	gtcaacacgg	gataataccg	cgccacatag	4680
cagaacttta	aaagtgctca	tcattggaaa	acgttcttcg	gggcgaaaac	tctcaaggat	4740
cttaccgctg	ttgagatcca	gttcgatgta	acccactcgt	gcacccaact	gatcttcagc	4800
atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaca	ggaaggcaaa	atgccgcaaa	4860
aaagggtaata	agggcgacac	ggaaatggtg	aatactcata	ctcttccttt	ttcaatatta	4920
ttgaagcatt	tatcaggggt	attgtctcat	gagcggatac	atatttgaat	gtatttagaa	4980
aaataaacia	ataggggttc	cgcgacacatt	tccccgaaaa	gtgccacctg	acgtctaaga	5040
aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggc	cctttcgtct	5100
tcaagaatte	tcatgtttga	cagcttatca	tcgataagct	ttaatgcggt	agtttatcac	5160
agttaaattg	ctaacgcagt	caggcaccgt	gtatgaaatc	taacaatgcg	ctcatcgtca	5220
tcctcggcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	gtactgccgg	5280
gcctcttgcg	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	gtgctgctag	5340
cgctatatgc	gttgatgcaa	tttctatgcg	cacccgttct	cggagcactg	tccgaccgct	5400
ttggccggcg	cccagtcctg	ctcgcttcgc	tacttggagc	cactatcgac	tacgggatca	5460
tggcgaccac	accgctcctg	tggatcccag	acgagttaag	tcaccatacg	ttagtacagg	5520
ttgccactct	tttggcagac	gcagacctac	ggctacaata	gcgaagcggg	cctgggtattc	5580
atgtttaaaa	atactgtcgc	gatagccaaa	acggcactct	ttggcagtta	agcgcacttg	5640
cttgccctgtc	gccagttcaa	cagaatcaac	ataagcgaac	actcgctgta	attctacgcc	5700
ataagcacca	atattctgga	taggtgatga	gccgacacaa	ccaggaatta	atgccagatt	5760
ttccagacca	ggcatacctt	cctgcaaagt	gtattttacc	agacgatgcc	agttttctcc	5820
ggctcctaca	tgtaaatacc	acgcatcagg	ttcatcatga	atttcgatac	ctttgatccg	5880
gttgatgatc	cctgcaggcc	cttaaggcca	tttaaatggc	gcgcccgatca	atgccaaatg	5940
tgttccaggg	ttttaaggag	tggttcatag	ctgctttcct	gatgcaaaaa	cgaggctagt	6000
ttaccgtatc	tgtgggggga	tggctttag	atatgacgac	aggaagagtt	tgtagaaacg	6060
caaaaaggcc	atccgtcagg	atggccttct	gcttaatttg	atgcctggca	gtttatggcg	6120
ggcgtcctgc	ccgccaccct	ccgggccggt	gcttcgcaac	gttcaaatcc	gctcccggcg	6180
gatttgtcct	actcaggaga	gcggttcaccg	acaaacaaca	gataaaacga	aaggcccagt	6240
ctttcgactg	agcctttcgt	tttatttgat	gcctggcagt	tcctactct	cgcatgggga	6300
gacccacac	taccatcggc	gctacggcgt	ttcacttctg	agttcggcat	ggggtcaggt	6360
gggaccaccg	cgctactgcc	gccaggcaaa	ttctgtttta	tcagaccgct	tctgcgttct	6420
gatttaatct	gtatcaggct	gaaaatcttc	tctcatccgc	caaaacagct	tcggcgttgt	6480
aaggttaagc	ctcacgggtc	attagtaccg	gttagctcaa	cgcatcgtct	cgcttacaca	6540
cccggcctat	caacgtcgtc	gtcttcaacg	ttccttcagg	acccttaaag	ggtcagggag	6600
aactcatctc	ggggcaagtt	tcgtgcttag	atgctttcag	cacttatctc	ttccgcattt	6660
agctaccggg	cagtgcatt	ggcatgacaa	cccgaacacc	agtgatgcgt	ccactccggt	6720
cctctcgtac	taggagcagc	ccccctcagt	tctccagcgc	ccacggcaga	tagggaccga	6780
actgtctcac	gacgttctaa	accagctcgc	cgtaccactt	taaatggcga	acagccatac	6840
ccttgggacc	tacttcagcc	ccaggatgtg	atgagccgac	atcgaggtgc	caaacaccgc	6900

-continued

---

cgtegatatg	aactcctggg	cggtatcagc	ctgttatccc	cggagtacct	tttatccggt	6960
gagcgatggc	ccttccatc	agaaccaccg	gatcactatg	acctgcttcc	gcacctgctc	7020
gcgccgtcac	gctcgcagtc	aagctggctt	atgccattgc	actaacctcc	tgatgtccga	7080
ccaggattag	ccaaccttcg	tgctcctcgg	ttactcttta	ggaggagacc	gccccagtca	7140
aactaccac	cagacactgt	cgcgaaccgg	gattacgggt	caacgttaga	acatcaaaca	7200
ttaaagggtg	gtatttcaag	gtcggctcca	tgacagctgg	cgtccacact	tcaaagcctc	7260
ccacctatcc	tacacatcaa	ggctcaatgt	tcagtgtcaa	gctatagtaa	aggttcacgg	7320
ggtctttccg	tcttgccgcg	ggtacactgc	atcttcacag	cgagttcaat	ttcactgagt	7380
ctcgggtgga	gacagcctgg	ccatcattac	gccattcgtg	caggtcggaa	cttaccggac	7440
aaggaatttc	gctaccttag	gaccgttata	gttacggccg	ccgtttaccg	gggcttcgat	7500
caagagcttc	gcttgccgta	accccatcaa	ttaaccttcc	ggcaccgggc	aggcgtcaca	7560
ccgtatacgt	ccactttcgt	gtttgcacag	tgctgtgttt	ttaataaaca	gttgcagcca	7620
gctggtatct	tcgactgatt	tcagctccat	ccgcgagggg	cctcacctac	atatcagcgt	7680
gccttctccc	gaagttacgg	caccatthttg	cctagttcct	tcacccgagt	tctctcaagc	7740
gccttggtat	tctctacctg	accacctgtg	tcggtttggg	gtacgatttg	atggtacctg	7800
atgcttagag	gcttttccctg	gaagcagggc	atgtgtgct	tcagcacctg	agtgcctcgt	7860
catcacgcct	cagccttgat	tttccggatt	tgcttgaaa	accagcctac	acgcttaaac	7920
cgggacaacc	gtcgcgccgg	caacatagcc	ttctccgtcc	ccccttcgca	gtaaacacaa	7980
gtacaggaat	attaacctgt	ttcccatcga	ctacgccttt	cggcctcggc	ttaggggtcg	8040
actcacctcg	ccccgattaa	cggtggacag	gaacccttgg	tcttcggcgg	agcgggcttt	8100
tcacccgctt	tatcgttact	tatgtcagca	ttcgcacttc	tgatacctcc	agcatgcctc	8160
acagcacacc	ttcgcagget	tacagaacgc	ttccctaccc	aacaacgcat	aagcgtcgtc	8220
gccgcagctt	cggtgcatgg	tttagccccg	ttacatcttc	cgcgagggcc	gactcgacca	8280
gtgagctatt	acgctttctt	taaatgatgg	ctgcttctaa	gccaacatcc	tggtgtctg	8340
ggccttccca	catcgtttcc	cacttaacca	tgactttggg	accttagctg	gcggctcggg	8400
ttgtttccct	cttcacgacg	gacgtagca	cccgcctgtg	gtctccctg	ataacattct	8460
ccggtattcg	cagtttgcac	cgggttggtg	agtcgggatg	accccttgc	cgaacagtg	8520
ctctaccccc	ggagatgaat	tcacgagggc	ctacctaaat	agctttcggg	gagaaccagc	8580
tatctcccgg	tttgattggc	ctttcacccc	cagccacaag	tcctccgcta	atthttcaac	8640
attagtcggg	tcggtcctcc	agttagtggt	acccaacctt	caacctgccc	atggctagat	8700
caccgggttt	cgggtctata	ccctgcaact	taacgcccag	ttaagactcg	gtttcccttc	8760
ggctccccta	ttcggttaac	cttgctacag	aatataagtc	gctgacctat	tatacaaaag	8820
gtacgcagtc	acacgcctaa	gcgtgctccc	actgcttgta	cgtaacaggt	ttcaggttct	8880
ttttcactcc	cctcgcgggg	gttcttttcg	cctttccctc	acggactcgg	ttcactatcg	8940
gtcagtcagg	agtatttagc	cttgagggat	ggcccccca	tattcagaca	ggataccacg	9000
tgtcccggcc	tactcatcga	gctcacagca	tgtgcatttt	tgtgtacggg	gctgtcacc	9060
tgtatcgcgc	gcctttccag	acgcttccac	taacacacac	actgattcag	gctctgggct	9120
gtccccggtt	cgctcggcgc	tactggggga	atctcgggtg	atthcttttc	ctcgggggtac	9180
ttagatgttt	cagttcccc	ggttcgcctc	attaacctat	ggattcagtt	aatgatagtg	9240

-continued

---

tgtcgaaaca	cactggggtt	cccattcgg	aaatcgccg	ttataacggt	tcatatcacc	9300
ttaccgacgc	ttatcgcaga	ttagcacgtc	cttcatcgcc	tctgactgcc	agggatcca	9360
ccgtgtacgc	ttagtcgctt	aacctcacia	cccgaagatg	tttctttcga	ttcatcatcg	9420
tgttgcgaaa	atitgagaga	ctcacgaaca	actctcggtg	ttcagtgttt	caatitctag	9480
cttgatccag	atititaaag	agcaaatatc	tcaaacatca	cccgaagatg	agtititgaga	9540
tattaaggtc	ggcgactttc	actcaciaac	cagcaagtgg	cgccccctag	gggattcgaa	9600
cccctgttac	cgccgtgaaa	ggggcggtgc	ctgggcctct	agacgaaggg	gacacgaaaa	9660
ttgcttatca	cgcgttgctg	gatatitctg	tgtagggtga	gctitcatta	atagaaagcg	9720
aacggcctta	ttctcttcag	cctcactccc	aacgcgtaaa	cgcttgcctt	ttcactitct	9780
atcagacaat	ctgtgtgagc	actacaaagt	acgcttcttt	aaggtaatcc	catgatccaa	9840
ccgcaggttc	ccctacggtt	accttgttac	gacttcaccc	cagtcatgaa	tactccgtg	9900
gtaaccgtcc	cccttgcggt	tagactagct	acttctggag	caaccactc	ccatgggtgtg	9960
acgggcggtg	tgtacaaggc	ccgggaacgt	atcaccgtg	acattctgat	tacgattac	10020
tagcgattcc	gacttcacgc	agtcgagttg	cagactgcga	tccggactac	gatcggtit	10080
atgggattag	ctccacctcg	cggttggtca	accctttgta	ccgaccattg	tagcacgtgt	10140
gtagccctgg	ccgtaagggc	catgatgact	tgacgtcacc	cccacttcc	tccggtitgt	10200
caccggcagt	ctccttagag	tgccaccctg	aggtgctggt	aactaaggac	aagggttgcg	10260
ctcgttacgg	gacttaacct	aacatctcac	gacacgagct	gacgacagcc	atgcagcacc	10320
tgtgtctgag	ttcccgaagg	caccaatcca	tctctggaaa	gttctcagca	tgtcaaggcc	10380
aggtaagggt	cttcgcttg	cttcgaatta	aaccacatgc	tccaccgctt	gtgcgggccc	10440
ccgtcaattc	atitgagttt	taaccttgcg	gccgtactcc	ccaggcggtc	gacttatcgc	10500
gtagctgctg	ccactaagat	ctcaaggatc	ccaaccgcta	gtcgacatcg	tttacggcgt	10560
ggactaccag	ggtatctaat	cctgtttgct	ccccacgctt	tgcacctca	gtgtcagtat	10620
cagtcacagt	ggtcgccttc	gccactgggtg	ttccttctca	tatctacgca	tttaccgctt	10680
acacaggaaa	ttccaccacc	ctctaccgta	ctctagctca	gtagtititgg	atgcagttcc	10740
caggttgagc	ccggggattt	cacatccaac	ttgtgaacc	acctacgcgc	gctttacgcc	10800
cagtaattcc	gattaacgct	tgcacccttc	gtattaccgc	ggctgctggc	acgaagttag	10860
ccggtgctta	ttctgttggt	aacgtcaaaa	cagcaaggta	ttaacttact	gcccttctct	10920
ccaacttaaa	gtgctttaca	atccgaagac	cttcttcaca	cacggggcat	ggctggatca	10980
ggctttcgcc	cattgtccaa	tattccccac	tgctgcctcc	cgtaggagtc	tggaccgtgt	11040
ctcagttcca	gtgtgactga	tcatctctct	agaccagtta	cggatcgtcg	ccttggtagg	11100
cctttacccc	accaactagc	taatccgacc	taggctcacc	tgatagcgtg	aggtccgaag	11160
atccccact	ttggtcttgc	gacgttatgc	ggtattagcg	cccgtttccg	gacgttatcc	11220
cccactacca	ggcagattcc	taggcattac	tcaccctgct	gccgctgaat	ccaggagcaa	11280
gtccccctca	tccgctcgac	ttgcatgtgt	taggcctgcc	gccagcgttc	aatctgagcc	11340
atgatcaaac	tcttcaattt	aaaagtttga	cgctcaaaaga	attaacttct	gtaatgaatt	11400
acgtgttcac	tcttgagact	tggattcat	tttctgtctt	gcgacgttaa	gaatccgtat	11460
cttcgagtg	ccacacagat	tgtctgataa	attgttaaag	agcagtgccg	cttcgctttt	11520
tctcagcggc	cgctgtgtga	aattgttatc	cgctcacaat	tccacacatt	atacgagccg	11580
gaagcataaa	gtgtaaagcc	tggggtgcct	aatgagtgag	ctaactcaca	ttaattgcgt	11640

-continued

---

```

tgcgctcact gcccgcttcc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg 11700
gccaacgcgc ggggagaggc gggttgcgta ttgggcgcca ggggtggttt tcttttcacc 11760
agtgagacgg gcaacagctg attgcccttc accgcctggc cctgagagag ttgcagcaag 11820
cggtcacgc tggtttccc cagcaggega aaatcctggt tgatggtggt tgacggcggg 11880
atataacatg agctgtcttc ggtatcgtcg tatcccacta ccgagatata cgcaccaacg 11940
cgcagcccgg actcggtaat ggcgcgcatt gcgcccagcg ccatctgata gttggcaacc 12000
agcatcgcag tgggaacgat gccctcattc agcatttgcg tggtttgttg aaaaccggac 12060
atggcactcc agtcgccttc ccgttccgct atcggctgaa tttgattgcg agtgagatat 12120
ttatgccagc cagccagacg cagacgcgcc gagacagaa ttaatgggcc cgctaacagc 12180
gcgatttgcg ggtgbcctaa tgcgaccaga tgctccacgc ccagtcgcgt accgtcttca 12240
tgggagaaaa taatactggt gatgggtgtc tggtcagaga catcaagaaa taacgccgga 12300
acattagtgc aggcagcttc cacagcaatg gcatcctggt catccagcgg atagttaata 12360
atcagcccac tgaccggtg cgcgagaaga ttgtgcaccg ccgctttaca ggcttcgacg 12420
ccgcttcggt ctaccatcga caccaccacg ctggcaccga gttgatcggc gcgagattta 12480
atcgcgcgca caatttgcga cggcgcgtgc agggccagac tggaggtggc aacgccaatc 12540
agcaacgact gtttgcggc cagttgttgt gccacgcggg tgggaatgta attcagctcc 12600
gccatgcgcg ctccacttt tcccgcggt ttcgcagaaa cgtggctggc ctggttcacc 12660
acgcgggaaa cggctctgata agagacaccg gcatactctg cgacatcgtg taacgttact 12720
ggtttcacat tcaccacct gaattgactc tctccgggc gctatcatgc cataccgcga 12780
aaggttttgc accattcgat ggtgtcggat cctctagcgg gaccacttg cggccacgat 12840
ccgtccgcgc taaggctcat accgttaatt attccccccc acgggagacc tgagcaaact 12900
gccctcaggc atttgagaag cacaggggtc cactgcttcg ggtagtcaat aaaccggtaa 12960
accagcaata gacataagcg gctatttaac gaccctgcc tgaaccgacg accgggtcga 13020
atttgcttcc gaatttctgc cattcatccg cttattatca cttattcagg cgtagcacca 13080
ggcgtttaag ggcaccaata actgccttaa aaaaa 13115

```

```

<210> SEQ ID NO 5
<211> LENGTH: 4459
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
nucleotide sequence

```

```

<400> SEQUENCE: 5
actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg 60
ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac 120
cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt 180
gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgctcgcag 240
caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta gcttcccggc 300
aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc 360
ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta 420
tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg 480
ggagtcaggc aactatgat gaacgaaata gacagatcgc tgagataggt gcctcactga 540

```

-continued

---

ttaagcattg	gtaactgtca	gaccaagttt	actcatatat	acttttagatt	gatttaaaac	600
ttcattttta	atttaaaagg	atctaggtga	agatcctttt	tgataatctc	atgacaaaa	660
tcccttaacg	tgagttttcg	ttccactgag	cgtcagaccc	cttaataaga	tgatcttctt	720
gagatcgttt	tggctctgcg	gtaatctctt	gctctgaaaa	cgaaaaaac	gccttgacgg	780
gcggtttttc	gaaggttctc	tgagctacca	actctttgaa	ccgaggtaac	tggcttgag	840
gagcgcagtc	acaaaaactt	gtcctttcag	tttagcctta	accggcgcat	gacttcaaga	900
ctaactcctc	taaatcaatt	accagtggct	gctgccagtg	gtgcttttgc	atgtctttcc	960
gggttgact	caagacgata	gttaccggat	aaggcgcagc	ggcggactg	aacggggggg	1020
tcgtgcatac	agtccagctt	ggagcgaact	gcctaccggg	aactgagtgt	caggcgtgga	1080
atgagacaaa	cgcgccata	acagcgaat	gacaccggtg	aaccgaaagg	caggaacagg	1140
agagcgcacg	agggagccgc	cagggggaaa	cgctgggtat	ctttatagtc	ctgtcggggt	1200
tcgccaccac	tgatttgagc	gtcagatttc	gtgatgcttg	tcaggggggc	ggagcctatg	1260
gaaaaacggc	tttgccgcg	ccctctcact	tcctgttaa	gtatcttctt	ggcatcttcc	1320
aggaaatctc	cgccccgttc	gtaagccatt	tcgctcgcc	gcagtcgaac	gaccgagcgt	1380
agcgagtcag	tgagcgagga	agcgggaat	atcctgtatc	acatattctg	ctgacgcacc	1440
ggtgcagcct	ttttctcct	gccacatgaa	gcacttcaact	gacaccctca	tcagtgccaa	1500
catagtaagc	cagtatacac	tcgctagcgc	ctgaggtctg	cctcgtgaag	aagggtgtgc	1560
tgactcatac	caggcctgaa	tcgccccatc	atccagccag	aaagtgaggg	agccacgggt	1620
gatgagagct	ttgtttagg	tggaccagtt	ggtgattttg	aacttttgct	ttgccacgga	1680
acggctctcg	ttgtcgggaa	gatgcgtgat	ctgatccttc	aactcagcaa	aagttcgatt	1740
tattcaacaa	agccacgttg	tgtctcaaaa	tctctgatgt	tacattgcac	aagataaaaa	1800
tatatcatca	tgaacaataa	aactgtctgc	ttacataaac	agtaatacaa	ggggtgttat	1860
gagccatatt	caacgggaaa	cgtcttgctc	gaggccgcga	ttaaattcca	acatggatgc	1920
tgatttatat	gggtataaat	gggctcgcga	taatgtcggg	caatcaggtg	cgacaatcta	1980
tcgattgtat	gggaagcccg	atgcgccaga	gttgtttctg	aaacatggca	aaggtagcgt	2040
tgccaatgat	gttacagatg	agatggtcag	actaaactgg	ctgacggaat	ttatgcctct	2100
tccgaccatc	aagcatttta	tcgctactcc	tgatgatgca	tggttactca	ccactgcat	2160
ccccgggaaa	acagcattcc	aggtattaga	agaatatcct	gattcaggtg	aaaatattgc	2220
tgatgcgctg	gcagtgttcc	tgcgccgggt	gcattcgatt	cctgtttgta	attgtccttt	2280
taacagcgat	cgcgatattc	gtctcgctca	ggcgcaatca	cgaatgaata	acggtttggg	2340
tgatgcgagt	gattttgatg	acgagcgtaa	tggctggcct	gttgaacaag	tctggaaaga	2400
aatgcataaa	cttttgccat	tctcaccgga	ttcagtcgtc	actcatgggt	atttctcact	2460
tgataacctt	atttttgacg	aggggaaatt	aataggttgt	attgatgttg	gacgagtcgg	2520
aatcgcagac	cgataccagg	atcttgccat	cctatggaac	tgctcgggtg	agttttctcc	2580
ttcattacag	aaacggcttt	ttcaaaaata	tggattgat	aatcctgata	tgaataaatt	2640
gcagtttcat	ttgatgctcg	atgagttttt	ctaatcagaa	ttggtaatt	ggttgtaaca	2700
ctggcagagc	attacgctga	cttgacggga	cggcgctttt	gttgaataaa	tcgaactttt	2760
gctgagttga	aggatcagat	cacgcactct	cccgacaacg	cagaccgttc	cgtggcaaag	2820
caaaagttca	aatcacttag	tcgacctagg	taccatcgat	gcataatgtg	cctgtcaaat	2880



-continued

---

```

ggacgaagca gggattctgc aaaccctatg ctactccgtc aagccgtcaa ttgtctgatt 2940
cgttaccaat tatgacaact tgacggctac atcattcact ttttcttcac aaccggcacg 3000
gaactcgctc gggctggccc cgggtgcattt tttaaatacc cgcgagaaat agagttgatc 3060
gtcaaaacca acattgacgac cgacgggtggc gatagggcatc cgggtggtgc tcaaaagcag 3120
cttcgectgg ctgatacggt ggtcctcgcg ccagcttaag acgctaatac ctaactgctg 3180
gcgaaaaaga tgtgacagac gcgacggcga caagcaaaca tgctgtgcca cgctggcgat 3240
atcaaaattg ctgtctgcca ggtgatcgct gatgtactga caagcctcgc gtacccgatt 3300
atccatcggt ggatggagcg actcgttaat cgcttccatg cgccgcagta acaattgctc 3360
aagcagattt atcgccagca gctccgaata gcgcccttcc ccttgcccgg cgtaaatgat 3420
ttgcccnaac aggtcgctga aatgctggctg gtgcgcttca tccgggagaa agaaccccg 3480
attggcaaat attgacggcc agttaagcca ttcattgccag taggcgcgcg gacgaaagta 3540
aaccactgg tgataccatt cgcgagcctc cggatgacga ccgtagtgat gaatctctcc 3600
tggcgggaac agcaaaatat caccggctcg gcaaacaaat tctcgtccct gatTTTTTcac 3660
caccctctga ccgcgaatgg tgagattgag aatataacct ttcattccca gcggtcggtc 3720
gataaaaaaa tcgagataac cgttggcctc aatcggcggt aaaccgcca ccagatgggc 3780
attaaacgag tatcccgcca gcaggggatc attttgcgct tcagccatac ttttcatact 3840
cccgccattc agagaagaaa ccaattgtcc atattgcac agacattgcc gtcactgctg 3900
cttttactgg ctcttctcgc taaccaaacc ggtaaccccg cttattaaaa gcattctgta 3960
acaaagcggg accaaagcca tgacaaaaac gcgtaacaaa agtgtctata atcacggcag 4020
aaaagtccac attgattatt tgcacggcgt cacactttgc tatgcatag catttttate 4080
cataagatta gcggatccta cctgacgctt tttatcgcaa ctctctactg tttctccatg 4140
cggccgcacg caagcttcta gaataaaacg aaaggctcag tcgaaagact gggcctttcg 4200
ttttatctgt tgtttgtcgg tgaacgctct cctgagtagg acaaatccgc cgggagcgga 4260
tttgaacggt gcgaagcaac ggcccggagg gtggcgggca ggacgcccgc cataaactgc 4320
caggcatcaa attaagcaga aggccatcct gacggatggc ctttttgaat tcaactggccg 4380
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgacg 4440
cacatcccc tttegccag 4459

```

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 4759

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Description of Artificial Sequence: Synthetic nucleotide sequence

&lt;400&gt; SEQUENCE: 6

```

actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg 60
ctgccataac catgagtgat aacctgccc ccaacttact tctgacaacg atcggaggac 120
cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcggt 180
gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgcag 240
caatggcaac aacgttgccg aaactattaa ctggcgaact acttactcta gcttcccggc 300
aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc 360
ttccgctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta 420

```

-continued

---

tcattgcagc	actggggcca	gatggtaagc	cctcccgtat	cgtagttatc	tacacgacgg	480
ggagtcaggc	aactatggat	gaacgaaata	gacagatcgc	tgagataggt	gcctcactga	540
ttaagcattg	gtaactgtca	gaccaagttt	actcatatat	acttttagatt	gatttaaaac	600
ttcattttta	atttaaaagg	atctaggtga	agatcctttt	tgataatctc	atgaccaaaa	660
tcccttaacg	tgagttttcg	ttccactgag	cgtcagaccc	cttaataaga	tgatcttctt	720
gagatcgttt	tggctcgcgc	gtaatctctt	gctctgaaaa	cgaaaaaac	gccttgcagg	780
gcggtttttc	gaaggttctc	tgagctacca	actctttgaa	ccgaggtaac	tggttggag	840
gagcgcagtc	acaaaaactt	gtcctttcag	tttagcctta	accggcgcac	gacttcaaga	900
ctaactcctc	taaatcaatt	accagtggct	gctgccagtg	gtgcttttgc	atgtctttcc	960
gggttggact	caagacgata	gttaccggat	aaggcgcagc	ggtcggactg	aacggggggg	1020
tcgtgcatac	agtccagctt	ggagcgaact	gcctaccogg	aactgagtgt	caggcgtgga	1080
atgagacaaa	cgccggcata	acagcggaat	gacaccggtg	aaccgaaagg	caggaacagg	1140
agagcgcacg	aggagccgc	cagggggaaa	cgctcgttat	ctttatagtc	ctgtcggggt	1200
tcgccaccac	tgatttgagc	gtcagatttc	gtgatgcttg	tcaggggggc	ggagcctatg	1260
gaaaaacggc	tttgccgcgg	ccctctcact	tccctgttaa	gtatcttctc	ggcatcttcc	1320
aggaaatctc	cgccccgttc	gtaagccatt	tccgctcgcc	gcagtcgaac	gaccgagcgt	1380
agcgagtcag	tgagcgagga	agcggaatat	atcctgtatc	acatattctg	ctgacgcacc	1440
ggtgcagcct	tttttctcct	gccacatgaa	gcacttcact	gacacctca	tcagtgccaa	1500
catagtaagc	cagtatacac	tccgctagcg	ctgaggtctg	cctcgtgaag	aagggtgtgc	1560
tgactcatac	caggcctgaa	tcgccccatc	atccagccag	aaagtgaggg	agccacgggt	1620
gatgagagct	ttgtttagg	tggaccagtt	ggtgattttg	aacttttgct	ttgccacgga	1680
acggtctgcg	ttgtcgggaa	gatgcgtgat	ctgatccttc	aactcagcaa	aagtctgatt	1740
tattcaacaa	agccacgttg	tgtctcaaaa	tctctgatgt	tacattgcac	aagataaaaa	1800
tatatcatca	tgaacaataa	aactgtctgc	ttacataaac	agtaatacaa	ggggtgttat	1860
gagccatatt	caacgggaaa	cgtcttgctc	gaggccgcga	ttaaattcca	acatggatgc	1920
tgatttatat	gggtataaat	gggctcgcga	taatgtcggg	caatcaggtg	cgacaatcta	1980
tcgattgtat	gggaagcccg	atgcgccaga	gttgtttctg	aaacatggca	aaggtagcgt	2040
tgccaatgat	gttacagatg	agatggtcag	actaaactgg	ctgacggaat	ttatgcctct	2100
tccgaccatc	aagcatttta	tccgtactcc	tgatgatgca	tggttactca	ccactgcgat	2160
ccccgggaaa	acagcattcc	aggtattaga	agaatatcct	gattcaggtg	aaaatattgc	2220
tgatgcgctg	gcagtgttcc	tgcgcgggtt	gcattcgatt	cctgtttgta	attgtccttt	2280
taacagcgat	cgcgtatttc	gtctcgtca	ggcgcaatca	cgaatgaata	acggtttggt	2340
tgatgcgagt	gattttgatg	acgagcgtaa	tggctggcct	gttgaacaag	tctggaaaga	2400
aatgcataaa	cttttgccat	tctcaccgga	ttcagtcgtc	actcatgggtg	atttctcact	2460
tgataacctt	atthttgacg	aggggaaatt	aataggttgt	attgatgttg	gacgagtcgg	2520
aatcgcagac	cgataccagg	atcttgccat	cctatggaac	tgctcgggtg	agttttctcc	2580
ttcattacag	aaacggcttt	ttcaaaaata	tggatttgat	aatcctgata	tgaataaatt	2640
gcagtttcat	ttgatgctcg	atgagttttt	ctaatcagaa	ttggttaatt	ggttgtaaca	2700
ctggcagagc	attacgctga	cttgacggga	cggcggcttt	gttgaataaa	tcgaactttt	2760
gctgagttga	aggatcagat	cacgcactct	cccgacaacg	cagaccgttc	cgtggcaaag	2820

-continued

---

```

caaaagttca aaatcactag tcgaccatgg taccatcgat gcataatgtg cctgtcaaat 2880
ggacgaagca gggattctgc aaaccctatg ctactccgtc aagccgtcaa ttgtctgatt 2940
cgttaccat tatgacaact tgacggctac atcattcact ttttcttcac aaccggcacg 3000
gaactcgctc gggctggccc cgggtgcattt tttaaatacc cgcgagaaat agagttgatc 3060
gtcaaaacca acattgcgac cgacgggtggc gatagggatc cgggtggtgc tcaaaagcag 3120
cttcgcctgg ctgatacgtt ggtcctcgcg ccagcttaag acgctaatac ctaactgctg 3180
gcggaaga tgtgacagac gcgacggcga caagcaaca tgctgtgca cgctggcgat 3240
atcaaaattg ctgtctgcca ggtgatcgct gatgtactga caagcctcgc gtacccgatt 3300
atccatcggg ggatggagcg actcgttaat cgcttccatg cgccgcagta acaattgctc 3360
aagcagattt atcgccagca gctccgaata gcgcccttc ccttgcccgg cgtaaatgat 3420
ttgcccacac aggtcgctga aatgcccgtg gtgcgcttca tccgggcaaga agaaccctgt 3480
attggcaaat attgacggcc agttaagcca ttcattgccg taggcgcgcg gacgaaagta 3540
aaccactgg tgataccatt cgcgagcctc cggatgacga ccgtagtgat gaatctctcc 3600
tggcgggaac agcaaaatat caccggctcg gcaaaaaat tctcgtccct gatttttcac 3660
caccctctga ccgcaatgg tgagattgag aatataacct ttcattcca gcggtcggtc 3720
gataaaaaaa tcgagataac cgttggcctc aatcggcgtt aaaccgcca ccagatgggc 3780
attaaacgag tatcccggca gcaggggatc attttgcgct tcagccatac tttcatact 3840
cccgcattc agagaagaaa ccaattgtcc atattgcac agacattgcc gtcactgcgt 3900
cttttactgg ctcttctcgc taaccaaacc ggtaaccccg cttattaata gcattctgta 3960
acaaagcggg accaaagcca tgacaaaaac gcgtaaaaa agtgtctata atcacggcag 4020
aaaagtccac attgattatt tgcacggcgt cacactttgc tatgcatag catttttatc 4080
cataagatta gcggatccta cctgacgctt tttatcgcaa ctctctactg tttctccatg 4140
cggccgccta tttaccagt ttttttcgag gagctcgacg atggccaaca caccttccgc 4200
caaaaaacgc gccaaacagg ctgagaagcg tcgacgccac aacgccagcc tgcgctccat 4260
ggtgcgcacc tacatcaaga acgctcgtgaa agccatcgac gccaaaggacc tggaaaaagc 4320
ccaggccgcc ttcaccgccc ctgtaccggg gatcgaccgc atggctgaca aaggcatcat 4380
ccacaagaac aaggctgctc gtcataagag ccgtctgagc ggccacatca aggcctcag 4440
caccgctgcc gcctaatac gaataaaacg aaaggctcag tcgaaagact gggcctttcg 4500
ttttatctgt tgtttgtcgg tgaacgctct cctgagtagg acaaatccgc cgggagcgga 4560
tttgaacggt gcgaagcaac ggcccggagg gtggcgggca ggacgcccgc cataaactgc 4620
caggcatcaa attaagcaga aggccatcct gacggatggc ctttttgaat tcaactggccg 4680
tcgttttaca acgctcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgag 4740
cacatcccc tttegccag 4759

```

```

<210> SEQ ID NO 7
<211> LENGTH: 11
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)..(7)
<223> OTHER INFORMATION: a, c, g, u, unknown or other

```

-continued

---

<400> SEQUENCE: 7

cannnnncuc g 11

<210> SEQ ID NO 8

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<220> FEATURE:

<221> NAME/KEY: modified\_base

<222> LOCATION: (4)..(8)

<223> OTHER INFORMATION: a, c, g, u, unknown or other

<400> SEQUENCE: 8

ucannnnnuu a 11

<210> SEQ ID NO 9

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 9

caggaggcuc g 11

<210> SEQ ID NO 10

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 10

ucaccuccuu a 11

<210> SEQ ID NO 11

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 11

cagugucguc g 11

<210> SEQ ID NO 12

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 12

ucacacacuu a 11

<210> SEQ ID NO 13

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

-continued

---

<220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 13

cauaucccuc g 11

<210> SEQ ID NO 14  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 14

ucagggauuu a 11

<210> SEQ ID NO 15  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 15

caaacaccuc g 11

<210> SEQ ID NO 16  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 16

ucaagagguu a 11

<210> SEQ ID NO 17  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 17

cauaccuc g 11

<210> SEQ ID NO 18  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 18

ucaugagguu a 11

<210> SEQ ID NO 19  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:

-continued

---

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 19

cauaauccuc g 11

<210> SEQ ID NO 20  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 20

ucagaggauu a 11

<210> SEQ ID NO 21  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 21

caaaauccuc g 11

<210> SEQ ID NO 22  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 22

ucaugagguu a 11

<210> SEQ ID NO 23  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 23

cacauaccuc g 11

<210> SEQ ID NO 24  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 24

ucaagagguu a 11

<210> SEQ ID NO 25  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

-continued

---

oligonucleotide

<400> SEQUENCE: 25

caccgaccuc g 11

<210> SEQ ID NO 26  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 26

ucaugggauu a 11

<210> SEQ ID NO 27  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 27

cauaucuccuc g 11

<210> SEQ ID NO 28  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 28

ucaugagguu a 11

<210> SEQ ID NO 29  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 29

caacuaccuc g 11

<210> SEQ ID NO 30  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 30

ucaagagguu a 11

<210> SEQ ID NO 31  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

-continued

---

<400> SEQUENCE: 31

cauauaccuc g	11
--------------	----

<210> SEQ ID NO 32

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 32

ucaagagguu a	11
--------------	----

<210> SEQ ID NO 33

<211> LENGTH: 18

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<220> FEATURE:

<221> NAME/KEY: modified\_base

<222> LOCATION: (10)..(13)

<223> OTHER INFORMATION: a, c, g, u, unknown or other

<400> SEQUENCE: 33

cauaucccun nnaaaug	18
--------------------	----

<210> SEQ ID NO 34

<211> LENGTH: 14

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 34

ggaucauggg auua	14
-----------------	----

<210> SEQ ID NO 35

<211> LENGTH: 18

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 35

cauaucccuc gagaaug	18
--------------------	----

<210> SEQ ID NO 36

<211> LENGTH: 14

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 36

ggaucauggg auua	14
-----------------	----

<210> SEQ ID NO 37

<211> LENGTH: 18

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence



-continued

---

<220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 37

cauaucuccuc gagaaug 18

<210> SEQ ID NO 38  
 <211> LENGTH: 14  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 38

ggaucaccuc cuua 14

<210> SEQ ID NO 39  
 <211> LENGTH: 18  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 39

cauaucuccuc cgcaaauug 18

<210> SEQ ID NO 40  
 <211> LENGTH: 14  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 40

ggaucauggg auua 14

<210> SEQ ID NO 41  
 <211> LENGTH: 18  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 41

cauaucuccuc cgcaaauug 18

<210> SEQ ID NO 42  
 <211> LENGTH: 14  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 42

ggaucaccuc cuua 14

<210> SEQ ID NO 43  
 <211> LENGTH: 18  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:

-continued

---

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 43

cauaucuccuc cugaaaug 18

<210> SEQ ID NO 44  
 <211> LENGTH: 14  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 44

ggaucauggg auua 14

<210> SEQ ID NO 45  
 <211> LENGTH: 18  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 45

cauaucuccuc ccaaaaug 18

<210> SEQ ID NO 46  
 <211> LENGTH: 14  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 46

ggaucauggg auua 14

<210> SEQ ID NO 47  
 <211> LENGTH: 18  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 47

cauaucuccuc cacaaaug 18

<210> SEQ ID NO 48  
 <211> LENGTH: 14  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 48

ggaucauggg auua 14

<210> SEQ ID NO 49  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

-continued

---

oligonucleotide  
 <220> FEATURE:  
 <221> NAME/KEY: modified\_base  
 <222> LOCATION: (3)..(7)  
 <223> OTHER INFORMATION: a, c, g, u, unknown or other

<400> SEQUENCE: 49

cannnnncuc g 11

<210> SEQ ID NO 50  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<220> FEATURE:  
 <221> NAME/KEY: modified\_base  
 <222> LOCATION: (4)..(8)  
 <223> OTHER INFORMATION: a, c, g, u, unknown or other

<400> SEQUENCE: 50

ucannnnnuu a 11

<210> SEQ ID NO 51  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 51

caggaggcuc g 11

<210> SEQ ID NO 52  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 52

ucaccuccuu a 11

<210> SEQ ID NO 53  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 53

caauccccuc g 11

<210> SEQ ID NO 54  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 54

ucaagggauu a 11

-continued

---

<210> SEQ ID NO 55  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 55

cauaccucuc g 11

<210> SEQ ID NO 56  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 56

ucaauggguu a 11

<210> SEQ ID NO 57  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 57

cacaguccuc g 11

<210> SEQ ID NO 58  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 58

ucagacgauu a 11

<210> SEQ ID NO 59  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 59

caaacgacuc g 11

<210> SEQ ID NO 60  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 60

ucagugauuu a 11

-continued

---

<210> SEQ ID NO 61  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 61

cauagggguc g 11

<210> SEQ ID NO 62  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 62

ucauuggguu a 11

<210> SEQ ID NO 63  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 63

caucuugcuc g 11

<210> SEQ ID NO 64  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 64

ucaggagguu a 11

<210> SEQ ID NO 65  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 65

caauuaucuc g 11

<210> SEQ ID NO 66  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 66

ucagaauuuu a 11

<210> SEQ ID NO 67

-continued

---

<211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 67  
  
 cacagaacug g 11

<210> SEQ ID NO 68  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 68  
  
 ucaaucaguu a 11

<210> SEQ ID NO 69  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 69  
  
 caaaguucuc g 11

<210> SEQ ID NO 70  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 70  
  
 ucaaugaguu a 11

<210> SEQ ID NO 71  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 71  
  
 caauucacuc g 11

<210> SEQ ID NO 72  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 72  
  
 ucagugaauu a 11

<210> SEQ ID NO 73  
 <211> LENGTH: 11

-continued

---

<212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 73

caacucacuc g 11

<210> SEQ ID NO 74  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 74

ucagaguguu a 11

<210> SEQ ID NO 75  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 75

caagggacuc g 11

<210> SEQ ID NO 76  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 76

ucaugggauu a 11

<210> SEQ ID NO 77  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 77

caugguucug g 11

<210> SEQ ID NO 78  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 78

ucaaagaguu a 11

<210> SEQ ID NO 79  
 <211> LENGTH: 11  
 <212> TYPE: RNA

-continued

---

<213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 79  
  
 cacaccacuc g 11  
  
 <210> SEQ ID NO 80  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 80  
  
 ucaugguuuu a 11  
  
 <210> SEQ ID NO 81  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 81  
  
 cacccaccuc g 11  
  
 <210> SEQ ID NO 82  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 82  
  
 ucaaaggguu a 11  
  
 <210> SEQ ID NO 83  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 83  
  
 caucccaguc g 11  
  
 <210> SEQ ID NO 84  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 84  
  
 ucaagggguu a 11  
  
 <210> SEQ ID NO 85  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence



-continued

---

<220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 85  
  
 caaacuccuc g 11

<210> SEQ ID NO 86  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 86  
  
 ucauacuaau a 11

<210> SEQ ID NO 87  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 87  
  
 cauagaucuc g 11

<210> SEQ ID NO 88  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 88  
  
 ucaagaguuu a 11

<210> SEQ ID NO 89  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 89  
  
 gaagugucuc g 11

<210> SEQ ID NO 90  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 90  
  
 ucaggagauu a 11

<210> SEQ ID NO 91  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:

-continued

---

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 91

caaauaucuc g 11

<210> SEQ ID NO 92  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 92

ucagagauuu a 11

<210> SEQ ID NO 93  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 93

cauaccucuc g 11

<210> SEQ ID NO 94  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 94

ucaugagguu a 11

<210> SEQ ID NO 95  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 95

cauaguacuc g 11

<210> SEQ ID NO 96  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 96

ucauggauuu a 11

<210> SEQ ID NO 97  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

-continued

---

oligonucleotide

<400> SEQUENCE: 97

caauccacug g 11

<210> SEQ ID NO 98  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 98

ucaguggauu a 11

<210> SEQ ID NO 99  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 99

cacagaucuc g 11

<210> SEQ ID NO 100  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 100

ucaggguuuu a 11

<210> SEQ ID NO 101  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 101

cauaggacuc g 11

<210> SEQ ID NO 102  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 102

ucaugcuauu a 11

<210> SEQ ID NO 103  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

-continued

---

<400> SEQUENCE: 103

caacuaacug g 11

<210> SEQ ID NO 104

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 104

ucauaguguu a 11

<210> SEQ ID NO 105

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 105

caaaauaucuc g 11

<210> SEQ ID NO 106

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 106

ucaagguauu a 11

<210> SEQ ID NO 107

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 107

caaaauaucuc g 11

<210> SEQ ID NO 108

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 108

ucaggagauu a 11

<210> SEQ ID NO 109

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

-continued

---

<400> SEQUENCE: 109

caguggucuc g	11
--------------	----

<210> SEQ ID NO 110

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 110

ucagaggauu a	11
--------------	----

<210> SEQ ID NO 111

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 111

cauauugguc g	11
--------------	----

<210> SEQ ID NO 112

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 112

ugauggaauu a	11
--------------	----

<210> SEQ ID NO 113

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 113

gaagcuacuc g	11
--------------	----

<210> SEQ ID NO 114

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 114

ucaggagauu a	11
--------------	----

<210> SEQ ID NO 115

<211> LENGTH: 11

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 115

-continued

---

caauggacuc g 11

<210> SEQ ID NO 116  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 116

ucaggagauu a 11

<210> SEQ ID NO 117  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 117

caaggcgcuc g 11

<210> SEQ ID NO 118  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 118

ucagaggguu a 11

<210> SEQ ID NO 119  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 119

caaacaucuc g 11

<210> SEQ ID NO 120  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 120

ucaagauguu a 11

<210> SEQ ID NO 121  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 121

-continued

---

caucccacuc g 11

<210> SEQ ID NO 122  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 122

ucagguauu a 11

<210> SEQ ID NO 123  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 123

cacugaucuc g 11

<210> SEQ ID NO 124  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 124

ucagaggauu a 11

<210> SEQ ID NO 125  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 125

cauauggcug g 11

<210> SEQ ID NO 126  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 126

ucagggauuu a 11

<210> SEQ ID NO 127  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 127

gaaagagcuc g 11

-continued

---

<210> SEQ ID NO 128  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 128  
  
 ugaagagguu a 11

<210> SEQ ID NO 129  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 129  
  
 caacgaacug g 11

<210> SEQ ID NO 130  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 130  
  
 ucagaguguu a 11

<210> SEQ ID NO 131  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 131  
  
 caucuaucuc g 11

<210> SEQ ID NO 132  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 132  
  
 ucaggagauu a 11

<210> SEQ ID NO 133  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide  
  
 <400> SEQUENCE: 133  
  
 cauaccucuc g 11



-continued

---

<210> SEQ ID NO 134  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 134

ucaugagguu a 11

<210> SEQ ID NO 135  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 135

cauauaacuc g 11

<210> SEQ ID NO 136  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 136

ucaagagauu a 11

<210> SEQ ID NO 137  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 137

caauagcuc g 11

<210> SEQ ID NO 138  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 138

ucaugagguu a 11

<210> SEQ ID NO 139  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 139

cacauagcuc g 11

-continued

---

<210> SEQ ID NO 140  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 140

ucaugagguu a 11

<210> SEQ ID NO 141  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 141

cagggaccuc g 11

<210> SEQ ID NO 142  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 142

ucaagagguu a 11

<210> SEQ ID NO 143  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 143

cauauccguc g 11

<210> SEQ ID NO 144  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 144

ucaugggguu a 11

<210> SEQ ID NO 145  
<211> LENGTH: 11  
<212> TYPE: RNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> SEQUENCE: 145

caacuaccug g 11

<210> SEQ ID NO 146

-continued

---

<211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 146

ucaugagguu a 11

<210> SEQ ID NO 147  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 147

cauauagcug g 11

<210> SEQ ID NO 148  
 <211> LENGTH: 11  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> SEQUENCE: 148

ucaagagguu a 11

<210> SEQ ID NO 149  
 <211> LENGTH: 1536  
 <212> TYPE: RNA  
 <213> ORGANISM: Pseudomonas aeruginosa

<400> SEQUENCE: 149

gaacugaaga guuugaucuau ggcucagauu gaacgcuggc ggcaggccua acacaugcaa 60  
 gucgagcggg ugaaggggagc uugcuccugg auucagcggc ggacggguga guaaugccua 120  
 ggaucugcc ugguaguggg ggauaacguc cggaaacggg cgcuaauacc gcuaucgucc 180  
 ugagggagaa aguggggggu cuucggaccu cacgcuauca gaugagccua ggucggauua 240  
 gcuaguuggu gggguaaagg ccuaccaagg cgacgauccg uaacuggucu gagaggauga 300  
 ucagucacac uggaacugag acacggucca gacuccuacg ggaggcagca guggggauua 360  
 uuggacaauug ggcgaaagcc ugauccagcc augcccgug ugugaagaag gucuucggau 420  
 uguaaagcac uuuaaguugg gaggaagggc aguaaguuaa uaccuugcug uuuugacguu 480  
 accaacagaa uaagcaccgg cuaacuucgu gccagcagcc gcgguauac gaagggugca 540  
 agcguaaauac ggaauuacug ggcguaaagc gcgcuaggu gguucagca guuggaugug 600  
 aaaucuccgg gcuaaccug ggaacugcau ccaaaacuac ugagcuagag uacgguagag 660  
 ggugguggaa uuuccugugu agcggugaaa ugcuagaua uaggaaggaa caccaguggc 720  
 gaaggcgacc accuggacug auacugacac ugaggugcga aagcgugggg agcaaacagg 780  
 auuagauacc cugguagucc acgccguaaa cgaugucgac uagccguugg gaucuuugag 840  
 aucuuagugg cgcagcuaac gcgauaaguc gaccgccugg ggaguacggc cgcaagguua 900  
 aaacucaaau gaauugacgg gggcccgcac aagcggugga gcaugugguu uaaucgaag 960  
 caacgcgaag aaccuuaccu ggccuugaca ugcugagaac uuuccagaga uggauuggug 1020

-continued

---

```

ccuucgggaa cucagacaca ggugcugcau ggcugucguc agcucguguc gugagauguu 1080
ggguuaaguc ccguaacgag cgcaaccuu guccuuaguu accagcaccu cgggugggca 1140
cucuaaggag acugccggug acaaaccgga ggaagguggg gaugacguca agucaucaug 1200
gcccuuacgg ccagggcuac acacgugcua caauggucgg uacaaaggu ugcgaagccg 1260
cgagguggag cuaauccau aaaaccgauc guaguccgga ugcgagucug caacucgacu 1320
gcgugaaguc ggaaucgcu guaaucguga aucagaaugu cacggugaau acguucccg 1380
gccuuguaca caccgcccg cacaccaugg gaguggguug cuccagaagu agcuagucua 1440
accgcaaggg ggacgguuac cacggaguga uucauggcug gggugaaguc guaacaaggu 1500
agccguaggg gaaccugcg cuggaucacc uccuaa 1536

```

&lt;210&gt; SEQ ID NO 150

&lt;211&gt; LENGTH: 1537

&lt;212&gt; TYPE: RNA

&lt;213&gt; ORGANISM: Escherichia coli

&lt;400&gt; SEQUENCE: 150

```

aaaugaaga guuugaucuu ggcucagauu gaacgcuggc ggcaggccua acacaugcaa 60
gucgagcggg ugaagggagc uugcuccugg auucagcggc ggacggguga guaaugucug 120
ggaaaucguc ugaugguggg ggauaacuac uggaaaaggu agcuaauacc gcauaacguc 180
gcaagaccaa agagggggac cuucggggccu cuugccaucg gaugugccca gaugggauua 240
gcuaguaggu gggguaacgg cucaccuagg cgacgaucac uagcuggucu gagaggaua 300
ccagccacac uggaacugag acacggucca gacuccuacg ggaggcagca guggggaaua 360
uugcacaauug ggcgcaagcc ugaugcagcc augcccgug uaugaagaag ggcuucggcu 420
uguaaaguac uuucagcggg gaggaagggg aguaaaguua auaccuuugc ucauuugacg 480
uuaccgcag aagaagcacc ggcuaacucc gugccagcag ccgcguaau acggagggug 540
caagcguuaa ucggaaauac ugggcuuaaa ggcacgcag gcgguuuguu aagucagaug 600
ugaaaucucc gggcucaacc ugggaacugc aucugauacu ggcaagcuug agucucguag 660
agggggguag aauuccaggu guagcgguga aaugcguaga gaucuggagg aaucgggug 720
gcaagggcg ccccgugac gaagacugac gcucaggugc gaaagcgug ggagcaaca 780
ggauuagaua cccugguagu ccacgccgua aacgaugucg acuuggaggu ugugcccuug 840
aggcguggcu uccggagcua acgcuuaag ucgaccgccu ggggaguacg gccgcaaggu 900
uaaaacucua augaaauugac gggggccccg acaagcggug gagcaugug uuuuuucga 960
ugcaacgcga agaaccuuac cuggucuuga cauccacgga aguuucaga gaugagaug 1020
ugccuucggg aaccgugaga caggcugca uggcugucgu cagcucgugu ugugaaaugu 1080
uggguuaagu cccgcaacga gcgcaaccu uauccuuugu ugccagcggu ccggccggga 1140
acucaaagga gacugccagu gauaaacugg aggaaggugg ggauagcuc aagucaucau 1200
ggccuuacg accagggcua cacacgugcu acaauggcg auacaaagag aagcgaccuc 1260
gagagagcaa gcgaccuca uaaagugcgu cguaguccg auuggagucu gcaacucgac 1320
uccaugaagu cggaucgcu aguaaucgug gaucagaug ccacgguga uacguucccg 1380
ggccuuguac acaccgccg ucacaccaug ggaguggguu gcaaaagaag uagguagcuu 1440
aaccuucggg agggcgcuua ccacuuugug auucauggcu ggggugaagu cguacaagg 1500
uaaccguagg ggaaccugcg guuggaucac cuccuaa 1537

```

-continued

---

<210> SEQ ID NO 151  
 <211> LENGTH: 1535  
 <212> TYPE: RNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 nucleotide sequence

<400> SEQUENCE: 151

gaacugaaga guuugaucu ggcucagau gaacgcuggc ggcaggccua acacaugcaa	60
gucgagcggg ugaaggggagc uugcuccugg auucagcggc ggacgggguga guaaugccua	120
ggaaucugcc ugguaguggg ggauaacguc cggaaacggg cgcuaauacc gcauaacguc	180
gcaagaccaa aguggggggu cuucggaccu cacgcuauca gaugagccua ggucggauua	240
gcuaguuggu gggguaaagg ccuaccaagg cgacgauccg uaacuggucu gagaggaua	300
ucagucacac uggaacugag acacggucca gacuccuacg ggaggcagca guggggaua	360
uuggacaaug ggcgaaaggc ugauccagcc augcccgug ugugaagaag gucuucggau	420
uguaaagcac uuuaguuggg aggaagggca guaaguuaa accuugcugu uuugacguua	480
ccaacagaau aagcaccggc uaacuucgug ccagcagccg cgguaauacg aaggugcaa	540
gcguaaucg gaauucugg gcguaaaggc cgcuaggug guucagcaag uuggauguga	600
aaucuccggg cucaaccugg gaacugcauc caaaacuacu gagcuagagu acgguagagg	660
gugguggauu uuccugugua gcguguaau gcguagauu aggaaggaac accaguggcg	720
aaggcgacca ccuggacuga uacugacacu gaggugcgaa agcgugggga gcaaacagga	780
uuagauacc ugguagucca cgccguaaac gaugucgacu agccguuggg auccuugaga	840
ucuuaguggc gcagcuaacg cgauaagucg accgccuggg gaguacggcc gcaagguua	900
aacucaaaug aaugacggg ggcgccaca agcgugggag caugugguuu aaucgaagc	960
aacgcgaaga accuuaccug gccuugacau gcugagaacu uuccagagau ggauuggugc	1020
cuucgggaac ucagacacag gugcugcaug gcugucguc gcuuguguc ugagauguug	1080
gguaagucc cguaacgagc gcaaccuug uccuuaguua ccagcaccuc gggugggac	1140
ucuaaggaga cugccgguga caaacgggag gaaggugggg augacgucaa gucaucaugg	1200
ccuuacggc caggguaca cacgugcuac aauggucggg acaaaggguu gcgaagccgc	1260
gaggugggagc uaaucacua aaaccgaucg uaguccggau cgcagucugc aacucgacug	1320
cgugaagucg gaucgcuag uaaucgugaa ucagauguc acggugaaua cguuccggg	1380
ccuugucac accgccguc acaccauggg aguggguugc uccagaagua gcuagucuaa	1440
ccgcaagggg gacgguuacc acggagugau ucauggcugg ggugaagucg uaacaaggua	1500
accguagggg aaccugcggu uggaucaccu ccua	1535

<210> SEQ ID NO 152  
 <211> LENGTH: 80  
 <212> TYPE: RNA  
 <213> ORGANISM: Pseudomonas aeruginosa

<400> SEQUENCE: 152

gugggggaua acguccgaa acgggcgua auaccgcaua cguccgagg gagaaagugg	60
gggaucucg gaccucagc	80

<210> SEQ ID NO 153  
 <211> LENGTH: 91  
 <212> TYPE: PRT  
 <213> ORGANISM: Pseudomonas aeruginosa

-continued

&lt;400&gt; SEQUENCE: 153

Met Ala Asn Thr Pro Ser Ala Lys Lys Arg Ala Lys Gln Ala Glu Lys  
 1 5 10 15  
 Arg Arg Ser His Asn Ala Ser Leu Arg Ser Met Val Arg Thr Tyr Ile  
 20 25 30  
 Lys Asn Val Val Lys Ala Ile Asp Ala Lys Asp Leu Glu Lys Ala Gln  
 35 40 45  
 Ala Ala Phe Thr Ala Ala Val Pro Val Ile Asp Arg Met Ala Asp Lys  
 50 55 60  
 Gly Ile Ile His Lys Asn Lys Ala Ala Arg His Lys Ser Arg Leu Ser  
 65 70 75 80  
 Gly His Ile Lys Ala Leu Ser Thr Ala Ala Ala  
 85 90

&lt;210&gt; SEQ ID NO 154

&lt;211&gt; LENGTH: 87

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Escherichia coli

&lt;400&gt; SEQUENCE: 154

Met Ala Asn Ile Lys Ser Ala Lys Lys Arg Ala Ile Gln Ser Glu Lys  
 1 5 10 15  
 Ala Arg Lys His Asn Ala Ser Arg Arg Ser Met Met Arg Thr Phe Ile  
 20 25 30  
 Lys Lys Val Tyr Ala Arg Ile Glu Ala Gly Asp Lys Ala Ala Ala Gln  
 35 40 45  
 Lys Ala Phe Asn Glu Met Gln Pro Ile Val Asp Arg Gln Ala Ala Lys  
 50 55 60  
 Gly Leu Ile His Lys Asn Lys Ala Ala Arg His Lys Ala Asn Leu Thr  
 65 70 75 80  
 Ala Gln Ile Asn Lys Leu Ala  
 85

&lt;210&gt; SEQ ID NO 155

&lt;211&gt; LENGTH: 106

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Thermus sp.

&lt;400&gt; SEQUENCE: 155

Met Ala Gln Lys Lys Pro Lys Arg Asn Leu Ser Ala Leu Lys Arg His  
 1 5 10 15  
 Arg Gln Ser Leu Lys Arg Arg Leu Arg Asn Lys Ala Lys Lys Ser Ala  
 20 25 30  
 Ile Lys Thr Leu Ser Lys Lys Ala Ile Gln Leu Ala Gln Glu Gly Lys  
 35 40 45  
 Ala Glu Glu Ala Leu Lys Ile Met Arg Lys Ala Glu Ser Leu Ile Asp  
 50 55 60  
 Lys Ala Ala Lys Gly Ser Thr Leu His Lys Asn Ala Ala Ala Arg Arg  
 65 70 75 80  
 Lys Ser Arg Leu Met Arg Lys Val Arg Gln Leu Leu Glu Ala Ala Gly  
 85 90 95  
 Ala Pro Leu Ile Gly Gly Gly Leu Ser Ala  
 100 105

I claim:

1. A method for identifying functional mutant ribosomes comprising:

- (a) transforming a set of host cells with a set of plasmids, each plasmid comprising a mutant *Pseudomonas aeruginosa* 16S rRNA gene and a selectable marker gene; wherein said mutant *Pseudomonas aeruginosa* 16S rRNA gene comprises at least one mutation, in addition to a mutant helix 9 sequence and a first mutant Anti-Shine-Dalgarno sequence; and said first selectable marker gene comprises a first mutant Shine-Dalgarno sequence; and wherein said first mutant Anti-Shine-Dalgarno sequence and said first mutant Shine-Dalgarno sequence are a mutually compatible pair; thereby forming a set of transformed host cells;
- (b) isolating from the set of transformed host cells those host cells which express the selectable marker gene product; and
- (c) sequencing the mutant *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (b), thereby identifying functional mutant ribosomes.

2. The method of claim 1, wherein the selectable marker gene is a green fluorescent protein gene.

3. The method of claim 1, wherein said mutant helix 9 sequence is SEQ ID NO: 1.

4. The method of claim 1, wherein the mutant Anti-Shine-Dalgarno sequence is selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146 and 148.

5. The method of claim 1, wherein the mutant Shine-Dalgarno sequence is selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145 and 147.

6. The method of claim 1, wherein the mutant Shine-Dalgarno sequence and the mutant Anti-Shine-Dalgarno sequence are a complementary pair selected from the group consisting of SEQ ID NOs: 7 and 8, 9 and 10, 11 and 12, 13 and 14, 15 and 16, 17 and 18, 19 and 20, 21 and 22, 23 and 24, 25 and 26, 27 and 28, 29 and 30, 31 and 32, 33 and 34, 35 and 36, 37 and 38, 39 and 40, 41 and 42, 43 and 44, 45 and 46, 47 and 48, 49 and 50, 51 and 52, 53 and 54, 55 and 56, 57 and 58, 59 and 60, 61 and 62, 63 and 64, 65 and 66, 67 and 68, 69 and 70, 71 and 72, 73 and 74, 75 and 76, 77 and 78, 79 and 80, 81 and 82, 83 and 84, 85 and 86, 87 and 88, 89 and 90, 91 and 92, 93 and 94, 95 and 96, 97 and 98, 99 and 100, 101 and 102, 103 and 104, 105 and 106, 107 and 108, 109 and 110, 111 and 112, 113 and 114, 115 and 116, 117 and 118, 119 and 120, 121 and 122, 123 and 124, 125 and 126, 127 and 128, 129 and 130, 131 and 132, 133 and 134, 135 and 136, 137 and 138, 139 and 140, 141 and 142, 143 and 144, 145 and 146, and 147 and 148.

7. The method of claim 1, wherein the functional mutant ribosomes of step (c) have regions of interest which comprise sequences of one or more nucleic acids which are conserved in each mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced; and the method further comprising:

(d) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (c) are mutated; and each mutant *Pseudomonas aeruginosa* 16S rRNA gene further comprises a mutant helix 9 sequence, and a second mutant Anti-Shine-Dalgarno sequence;

(e) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (d) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second selectable marker having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;

(f) transforming a second set of host cells with the plasmids from step (e), thereby forming a second set of transformed host cells;

(g) isolating from the second set of transformed host cells from step (f) those host cells which express the selectable marker gene product; and

(h) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (g), thereby identifying functional mutant ribosomes that may be suitable as drug targets.

8. The method of claim 7, wherein the selectable marker gene is a green fluorescent protein gene, and the second selectable marker is a green fluorescent protein.

9. The method of claim 7, wherein said mutant helix 9 sequence is SEQ ID NO: 1.

10. The method of claim 7, wherein the mutant Anti-Shine-Dalgarno sequence and the second mutant Anti-Shine-Dalgarno sequence are selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146 and 148.

11. The method of claim 7, wherein the mutant Shine-Dalgarno sequence and the second mutant Shine-Dalgarno sequence are each selected from the group consisting of SEQ ID NOs: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145 and 147.

12. The method of claim 7, wherein the mutant Shine-Dalgarno sequence and the mutant Anti-Shine-Dalgarno sequence and the second mutant Anti-Shine-Dalgarno sequence are each a complementary pair selected from the group consisting of SEQ ID NOs: 7 and 8, 9 and 10, 11 and 12, 13 and 14, 15 and 16, 17 and 18, 19 and 20, 21 and 22, 23 and 24, 25 and 26, 27 and 28, 29 and 30, 31 and 32, 33 and 34, 35 and 36, 37 and 38, 39 and 40, 41 and 42, 43 and 44, 45 and 46, 47 and 48, 49 and 50, 51 and 52, 53 and 54, 55 and 56, 57 and 58, 59 and 60, 61 and 62, 63 and 64, 65 and 66, 67 and 68, 69 and 70, 71 and 72, 73 and 74, 75 and 76, 77 and 78, 79 and 80, 81 and 82, 83 and 84, 85 and 86, 87 and 88, 89 and 90, 91 and 92, 93 and 94, 95 and 96, 97 and 98, 99 and 100, 101 and 102, 103 and 104, 105 and 106, 107 and 108, 109 and 110, 111 and 112, 113 and 114, 115 and 116, 117 and 118, 119 and 120, 121 and 122, 123 and 124, 125 and 126, 127 and 128, 129 and 130, 131 and 132, 133 and

134, 135 and 136, 137 and 138, 139 and 140, 141 and 142, 143 and 144, 145 and 146, and 147 and 148.

**13.** The method of claim **1**, wherein the functional mutant ribosomes of step (c) have regions of interest which comprise sequences of one or more nucleic acids which are conserved in each mutant *Pseudomonas aeruginosa* 16S rRNA gene sequenced; and the method further comprising:

(d) generating a second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes wherein the regions of interest from step (c) are mutated; and each mutant *Pseudomonas aeruginosa* 16S rRNA gene further comprises a mutant helix 9 sequence, and a second mutant Anti-Shine-Dalgarno sequence;

(e) inserting the second set of mutant *Pseudomonas aeruginosa* 16S rRNA genes comprising the mutated regions of interest from step (d) into a second set of plasmids; wherein said plasmids further comprise a second genetically engineered gene which encodes a second selectable marker having a second mutant Shine-Dalgarno sequence, wherein the second mutant Anti-Shine-Dalgarno and the second mutant Shine-Dalgarno sequence are a mutually compatible pair;

(f) transforming a second set of host cells with the plasmids from step (e), thereby forming a second set of transformed host cells;

(g) isolating from the second set of transformed host cells from step (f) those host cells which express the selectable marker gene product;

(h) sequencing the *Pseudomonas aeruginosa* 16S rRNA gene from each host cell isolated in step (g), to identify the mutated regions of interest;

(i) screening compounds against the mutated regions of interest from step (h) and wildtype *Pseudomonas aeruginosa* 16S rRNA;

(j) identifying the compounds from step (i) that bind to the mutated regions of interest from step (h) and the wildtype *Pseudomonas aeruginosa* 16S rRNA;

(k) screening the compounds from step (j) against human 16S rRNA; and

(l) identifying the drug candidates from step (k) that do not bind to the human 16S rRNA, thereby identifying drug candidates.

**14.** The method of claim **13**, wherein the selectable marker gene is a green fluorescent protein gene, and the second selectable marker is a green fluorescent protein.

**15.** The method of claim **13**, wherein said mutant helix 9 sequence is SEQ ID NO: 1.

**16.** The method of claim **13**, wherein the mutant Anti-Shine-Dalgarno sequence and the second mutant Anti-Shine-Dalgarno sequence are selected from the group consisting of SEQ ID NOS: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146 and 148.

**17.** The method of claim **13**, wherein the mutant Shine-Dalgarno sequence and the second mutant Shine-Dalgarno sequence are each selected from the group consisting of SEQ ID NOS: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145 and 147.

**18.** The method of claim **13**, wherein the mutant Shine-Dalgarno sequence and the mutant Anti-Shine-Dalgarno sequence and the second mutant Shine-Dalgarno sequence and the second mutant Anti-Shine-Dalgarno sequence are each a complementary pair selected from the group consisting of SEQ ID NOS: 7 and 8, 9 and 10, 11 and 12, 13 and 14, 15 and 16, 17 and 18, 19 and 20, 21 and 22, 23 and 24, 25 and 26, 27 and 28, 29 and 30, 31 and 32, 33 and 34, 35 and 36, 37 and 38, 39 and 40, 41 and 42, 43 and 44, 45 and 46, 47 and 48, 49 and 50, 51 and 52, 53 and 54, 55 and 56, 57 and 58, 59 and 60, 61 and 62, 63 and 64, 65 and 66, 67 and 68, 69 and 70, 71 and 72, 73 and 74, 75 and 76, 77 and 78, 79 and 80, 81 and 82, 83 and 84, 85 and 86, 87 and 88, 89 and 90, 91 and 92, 93 and 94, 95 and 96, 97 and 98, 99 and 100, 101 and 102, 103 and 104, 105 and 106, 107 and 108, 109 and 110, 111 and 112, 113 and 114, 115 and 116, 117 and 118, 119 and 120, 121 and 122, 123 and 124, 125 and 126, 127 and 128, 129 and 130, 131 and 132, 133 and 134, 135 and 136, 137 and 138, 139 and 140, 141 and 142, 143 and 144, 145 and 146, and 147 and 148.

\* \* \* \* \*



UNITED STATES PATENT AND TRADEMARK OFFICE  
**CERTIFICATE OF CORRECTION**

PATENT NO. : 9,677,110 B2  
APPLICATION NO. : 14/746043  
DATED : June 13, 2017  
INVENTOR(S) : Cunningham

Page 1 of 1

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

In the Specification

Column 1, Line 22: "NIH SBIR Grant Number AI060275-01 awarded by the" should read  
-- AI060275 awarded by --.

Signed and Sealed this  
First Day of August, 2017



Joseph Matal  
*Performing the Functions and Duties of the  
Under Secretary of Commerce for Intellectual Property and  
Director of the United States Patent and Trademark Office*